

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 14:36:44 ; Search time 49.5417 Seconds
(without alignments)
828.417 Million cell updates/sec

Title: US-10-087-013-11

Perfect score: 1706

Sequence: 1 GNGGSRNISCNPKRESYPDW.....KKPEYNSKKAASEYLKDK 308

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

al number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 1706 | 100.0 | 308 | 22 | AAAB62151 |
| 2 | 894 | 52.4 | 311 | 22 | AAAB62150 |
| 3 | 669.5 | 39.2 | 2703 | 16 | AAAB62150 |
| 4 | 669.5 | 39.2 | 2710 | 18 | AAAB62150 |
| 5 | 669.5 | 39.2 | 2710 | 21 | AAAB62150 |
| 6 | 669.5 | 39.2 | 3060 | 18 | AAAB62150 |
| 7 | 669.5 | 39.2 | 3060 | 21 | AAAB62150 |
| 8 | 600 | 35.2 | 407 | 22 | AAAB62148 |
| 9 | 572 | 33.5 | 3542 | 22 | AAAB62142 |
| 10 | 521 | 30.5 | 793 | 16 | AAAB62142 |

| | | | | | |
|----|-------|------|------|----|-----------|
| 11 | 521 | 30.5 | 921 | 18 | AAAB62148 |
| 12 | 521 | 30.5 | 921 | 21 | AAAB62147 |
| 13 | 514 | 30.1 | 294 | 22 | AAAB62147 |
| 14 | 444 | 26.0 | 351 | 22 | AAAB62149 |
| 15 | 386 | 22.6 | 2228 | 20 | AAAB62149 |
| 16 | 367.5 | 21.5 | 2182 | 18 | AAAB62146 |
| 17 | 367.5 | 21.5 | 2182 | 21 | AAAB62146 |
| 18 | 342.5 | 20.1 | 2197 | 21 | AAAB62146 |
| 19 | 327.5 | 19.2 | 1143 | 23 | AAAB62146 |
| 20 | 310 | 18.2 | 2913 | 17 | AAAB62146 |
| 21 | 296 | 17.4 | 1726 | 17 | AAAB62146 |
| 22 | 268.5 | 15.7 | 749 | 16 | AAAB62146 |
| 23 | 268.5 | 15.7 | 749 | 18 | AAAB62146 |
| 24 | 268.5 | 15.7 | 749 | 21 | AAAB62146 |
| 25 | 268.5 | 15.7 | 1086 | 23 | AAAB62146 |
| 26 | 263.5 | 15.4 | 311 | 21 | AAAB62146 |
| 27 | 262.5 | 15.4 | 972 | 23 | AAAB62146 |
| 28 | 261 | 15.3 | 1421 | 23 | AAAB62146 |
| 29 | 260 | 15.2 | 616 | 23 | AAAB62146 |
| 30 | 259 | 15.2 | 1028 | 14 | AAAB62146 |
| 31 | 259 | 15.2 | 1115 | 12 | AAAB62146 |
| 32 | 259 | 15.2 | 1115 | 18 | AAAB62146 |
| 33 | 258.5 | 15.2 | 1115 | 21 | AAAB62146 |
| 34 | 258.5 | 15.2 | 1501 | 23 | AAAB62146 |
| 35 | 256.5 | 15.0 | 329 | 23 | AAAB62146 |
| 36 | 252 | 14.8 | 1061 | 16 | AAAB62146 |
| 37 | 251 | 14.7 | 350 | 23 | AAAB62146 |
| 38 | 251 | 14.7 | 1435 | 16 | AAAB62146 |
| 39 | 251 | 14.7 | 1435 | 18 | AAAB62146 |
| 40 | 251 | 14.7 | 1435 | 21 | AAAB62146 |
| 41 | 251 | 14.7 | 1604 | 16 | AAAB62146 |
| 42 | 251 | 14.7 | 1786 | 14 | AAAB62146 |
| 43 | 248.5 | 14.6 | 1245 | 16 | AAAB62146 |
| 44 | 248.5 | 14.6 | 1807 | 22 | AAAB62146 |
| 45 | 248.5 | 14.6 | 2028 | 22 | AAAB62146 |

ALIGNMENTS

| | | |
|----------|---|----------------------------|
| RESULT 1 | AAAB62151 | standard; peptide: 308 AA. |
| ID | AAAB62151 | standard; peptide: 308 AA. |
| AC | AAAB62151 | |
| DT | 29-MAY-2001 | (first entry) |
| XX | P. falciparum varCSA polypeptide Itg2-CS2 DBL2. | |
| DE | FCR3, varCSA protein; chondroitin sulfate A; CSA; var gene; PFEMP1; | |
| KW | erythrocyte membrane protein 1; parasitized red blood cell; PPRC; | |
| KW | malaria; protozoacide; Itg2-CS2 DBL2... | |
| XX | Plasmodium falciparum. | |
| OS | Plasmodium falciparum. | |
| XX | WO200116326-A2. | |
| XX | 08-MAR-2001. | |
| XX | 01-SEP-2000; 2000WO-US24195. | |
| XX | 01-SEP-1999; 99US-0152023. | |
| XX | (USSH) US DEPT HEALTH & HUMAN SERVICES. | |
| XX | Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C; | |
| XX | Gysin J, Pouvelle B, Fujii N, Smith J; | |
| XX | WPL: 2001-235109/24. | |
| XX | Novel FCR3, varCSA protein, useful for modulating parasitized red blood | |
| XX | cell binding, sequestration and onset of maternal malaria - | |

XX PS Claim 54; Page 75-76; 78pp; English.

CC The invention relates to a P. falciparum FCRI3 varCSA protein, that is capable of binding to chondroitin sulfate A (CSA). The var gene and the corresponding P.falciparum erythrocyte membrane protein 1 (PFEMP1) module adhesion of parasitized red blood cell (PRBC) to CSA. The protein and the encoding gene are useful for treating and preventing maternal malaria or in a patient afflicted with maternal malaria. The present sequence represents a P. falciparum varCSA polypeptide.

CC Itc2-CS2 DBL2.

CC

XX Sequence 308 AA;

SO

Query Match 100.0%; Score 1706; DB 22; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.4e-145;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GNDSSNETSGCNPKESYPDMDCCKNNIDNSHSGACMPRRROKLCVRDLTGGGEIRKPEDIL 60
1 GNDSSNETSGCNPKESYPDMDCCKNNIDNSHSGACMPRRROKLCVRDLTGGGEIRKPEDIL 60

QY 61 TKFINCAKETHFAMHKKYKKNVNAENELKSGKIPGFRKQMYTTFGDFRDIFFGTDISS 120
61 TKFINCAKETHFAMHKKYKKNVNAENELKSGKIPGFRKQMYTTFGDFRDIFFGTDISS 120

QY 121 CRYIKDTSQITKSLGDOATEKGDTHIDNKKLOEMWTTHGPKIWEGLCALTNGLSES 180
121 CRYIKDTSQITKSLGDOATEKGDTHIDNKKLOEMWTTHGPKIWEGLCALTNGLSES 180

DB 121 CRYIKDTSQITKSLGDOATEKGDTHIDNKKLOEMWTTHGPKIWEGLCALTNGLSES 180

QY 181 EKKNIIDYSYNKLNAEKDDCCELEKFAKPOPLRMVYEWSDFCRERKKLEDEVEDVCI 240
181 EKKNIIDYSYNKLNAEKDDCCELEKFAKPOPLRMVYEWSDFCRERKKLEDEVEDVCI 240

DB 181 EKKNIIDYSYNKLNAEKDDCCELEKFAKPOPLRMVYEWSDFCRERKKLEDEVEDVCI 240

QY 241 KAKYIEGCKNNKSNNSCVKCKEYENITGKKTQYEOEGFNTERKKQKREYNSYSKD 300
241 KAKYIEGCKNNKSNNSCVKCKEYENITGKKTQYEOEGFNTERKKQKREYNSYSKD 300

DB 241 KAKYIEGCKNNKSNNSCVKCKEYENITGKKTQYEOEGFNTERKKQKREYNSYSKD 300

QY 301 ASEYLKDK 308
301 ASEYLKDK 308

DB 301 ASEYLKDK 308

RESULT 2
AAB62150
ID AAB62150 standard; peptide; 311 AA.

XX AAB62150;

XX 29-MAY-2001 (first entry)

XX P. falciparum varCSA polypeptide FCRI3 var3DBL-gamma.

XX DE

XX FCRI3 varCSA protein; chondroitin sulfate A; CSA; var gene; PFEMP1; erythrocyte membrane protein 1; parasitized red blood cell; PRBC; malaria; protozoacide; FCRI3 var3DBL-gamma.

XX OS Plasmodium falciparum.

XX PN WO200116326-A2.

XX PD 08-MAR-2001.

XX PF 01-SEP-2000; 2000WO-US24195.

XX PR 01-SEP-1999; 99US-0152023.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PA Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C, Gysin J, Pouvelle B, Fujii N, Smith J;

XX PI

XX

DR WPI; 2001-235109/24.

XX

PT Novel FCRI3 varCSA protein, useful for modulating parasitized red blood cell binding, sequestration and onset of maternal malaria

XX

PS Disclosure; Page 74-75; 78pp; English.

XX

CC The invention relates to a P. falciparum FCRI3 varCSA protein, that is capable of binding to chondroitin sulfate A (CSA). The var gene and the corresponding P.falciparum erythrocyte membrane protein 1 (PFEMP1) module adhesion of parasitized red blood cell (PRBC) to CSA. The protein and the encoding gene are useful for treating and preventing maternal malaria or in a patient afflicted with maternal malaria. The present sequence represents a P. falciparum varCSA polypeptide.

CC FCRI3 var3DBL-gamma.

CC

XX Sequence 311 AA;

SO

Query Match 52.4%; Score 894; DB 22; Length 311;
Best Local Similarity 55.0%; Pred. No. 2.6e-72;
Matches 172; Conservative 36; Mismatches 91; Indels 12; Gaps 4;

5 SNETSGCNP-----KESYPDMDCCKNNIDNSHSGACMPRRROKLCVRDLTGGGEIRKPEDIL 59
5 SNETSGCNP-----KESYPDMDCCKNNIDNSHSGACMPRRROKLCVRDLTGGGEIRKPEDIL 59

DB 2 TDDIDGCKNOKYKAGKADKXPGWDCNSQIHTHNGACMPRRROKLCVSLGTTRDKAIKLEYI 61
2 TDDIDGCKNOKYKAGKADKXPGWDCNSQIHTHNGACMPRRROKLCVSLGTTRDKAIKLEYI 61

QY 60 LTKFINCAKETHFAMHKKYKKNVNAENELKSGKIPGFRKQMYTTFGDFRDIFFGTDISS 119
60 LTKFINCAKETHFAMHKKYKKNVNAENELKSGKIPGFRKQMYTTFGDFRDIFFGTDISS 119

DB 62 RTEFIKSAALIEHFAWRYKEDNGEAELEKNGNIPGFRKQMYTTFGDFRDIFFGTDISS 121
62 RTEFIKSAALIEHFAWRYKEDNGEAELEKNGNIPGFRKQMYTTFGDFRDIFFGTDISS 121

QY 120 SCRYIKDTSO--TIKSLGDOATEKGDTHIDNKKLOEMWTTHGPKIWEGLCALTNGL 176
120 SCRYIKDTSO--TIKSLGDOATEKGDTHIDNKKLOEMWTTHGPKIWEGLCALTNGL 176

DB 122 THAIVISGSKVITIIERKENDAKYAAKONS--NNELDDWMDHGXNDIBGMLCALTHK 178
122 THAIVISGSKVITIIERKENDAKYAAKONS--NNELDDWMDHGXNDIBGMLCALTHK 178

QY 177 LSESE--KKNIIDYSYNKLNAEKDDCCELEKFAKPOPLRMVYEWSDFCRERKKLEDEKY 235
177 LSESE--KKNIIDYSYNKLNAEKDDCCELEKFAKPOPLRMVYEWSDFCRERKKLEDEKY 235

DB 179 ISDEKKRKEIKNKYSYKRLNESPGSNKVEDFAKPOPLRMVYEWSDFCRERKKLEDEKY 238
179 ISDEKKRKEIKNKYSYKRLNESPGSNKVEDFAKPOPLRMVYEWSDFCRERKKLEDEKY 238

QY 236 EDVCIRKADYEGCKNNKSNNSCVKCKEYENITGKKTQYEOEGFNTERKKQKREYNSY 255
236 EDVCIRKADYEGCKNNKSNNSCVKCKEYENITGKKTQYEOEGFNTERKKQKREYNSY 255

DB 239 KVSQSDAKDYDGCCKNTSNASCVSACKVYEDYITKKVETKOKGKFDKAITDKSEYEG 298
239 KVSQSDAKDYDGCCKNTSNASCVSACKVYEDYITKKVETKOKGKFDKAITDKSEYEG 298

QY 296 YSKRDASEYLKDK 308
296 YSKRDASEYLKDK 308

DB 299 FSTKDASEYLKDK 311
299 FSTKDASEYLKDK 311

RESULT 3
AAR70236
ID AAR70236 standard; Protein; 2703 AA.

XX AAR70236;

XX AC AAR70236;

XX

XX 22-SEP-1995 (first entry)

XX DE P. falciparum Proj3.

XX DE

XX Erythrocyte binding ligand; Proj3; binding domain; malaria; therapy; vaccine.

XX KW

XX Plasmodium falciparum.

XX OS

XX PN WO9507353-A.

XX PD 16-MAR-1995.

XX PF 07-SEP-1994; 94WO-US10230.

XX PR 10-SEP-1993; 93US-0119677.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PA

XX

AC AAY77904;
 XX
 DT 13-JUN-2000 (first entry)
 XX
 DE P. falciparum Proj3 binding domain polypeptide.
 XX
 KW DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
 KM DABP; Stialic Acid Binding Protein; SABP; malaria; vaccine; immunisation;
 KM protozoacide; Proj3.
 XX
 OS Plasmodium falciparum.
 XX
 PN US5993827-A.
 XX
 PD 30-NOV-1999.
 XX
 PF 07-JUN-1995; 95US-0487826.
 XX
 PR 10-SEP-1993; 93US-0119677.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Sim KL, Chitnis C, Peterson DS, Su X, Wellems TE, Miller LH;
 XX
 DR WPI: 2000-194198/17.
 DR N-PSDB; AAY98286.

PT Isolated protein binding domains from Plasmodium vivax and Plasmodium
 PT falciparum erythrocyte binding proteins useful for vaccinating against
 PT malaria -
 XX
 PS Disclosure; Columns 79-92; 93pp; English.

XX The invention relates to ebl-1 polypeptides that are encoded by the DBL
 CC (Duffy-binding like) gene family. The ebl-1 proteins are substantially
 CC identical to the Duffy Antigen Binding Protein (DABP) and Stialic Acid
 CC Binding Protein (SABP), which are soluble proteins that appear in the
 CC culture supernatant after erythrocytes infected with malaria release
 CC merozoites. Immunochemical studies indicate that DABP and SABP are the
 CC respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy
 CC and stialic acid receptors on erythrocytes. The ebl-1 polypeptides may be
 CC used to vaccinate against malaria, especially caused by P. falciparum.
 CC Immunization with the polypeptide provides effective protection against
 CC malaria. The present sequence represents the Proj3 binding domain
 CC polypeptide.
 CC
 SO Sequence 2710 AA;

Query Match 39.2%; Score 669.5; DB 21; Length 2710;
 Best Local Similarity 40.9%; Pred. No. 8.9e-51;
 Matches 133; Conservative 54; Mismatches 109; Indels 29; Gaps 9;

QY 1 GNDGSENEISGNCNPKESYDPMDCKKNIDNSHGACMPRRKLC---VRDLTGGGELRKPE 57
 DB 1358 GKNQRTVGVGCNPKESYDPMDCKKNIDNSHGACMPRRKLC---VRLTGGGELRKPE 57
 QY 58 DITLFLNCACAKENHFMKYYKKDNVAENELSGKPEFRKQMTYTFDFDFDIFGCD 117
 DB 1418 NLMDATLKTAAEFETFSWQYKSKNSSEAKILDRGLPSQFLRSMTYTFDFDYDIDCLND 1477
 QY 118 ISSCRYIKDTSQTIKSLG-----DOATTEKGDTHIDNNKLEDMWTTHPKWMEGLCA 172
 DB 1478 ISKKQ--NDVAKA-KKIKGFESKDSKSPSGLSR-----QEMWKTNPETLWKGMLCA 1527
 QY 173 LTNGLSESE-KKNIIDYSTINKLNNAEKDDCCLEKFAKROFLRMVYVENSDFCRERKRL 231
 DB 1528 LTNVYVTDLTKRKIKNDYSIDKYNQSNPNLEEFARQPOFLRMVMEGEDEFCAROKK 1587
 QY 232 EDKVEVCIKAKYIECKKNKSNNGCVKCKEYENITGKTQYIEQEGKF--NTKKRK 289
 DB 1588 ENIKACMEINSTQOC--NDAKHRCNQACRAQVEYENKKEFGSGTNNFVLKANVQPO 1645
 QY 290 KPEYNSYSKKD-----ASEYLKDK 308

DB 1646 DPEYKGYEYKGVOPIOGNEYLLQK 1670

RESULT 6
 ID AAW22475
 XX AAW22475 standard; Protein: 3060 AA.
 AC
 XX AAW22475;
 DT 12-SEP-1997 (first entry)
 XX
 DE Plasmodium var-7.
 XX
 KW DBL gene family; SABP; stialic acid binding protein; vaccine; therapy;
 KM Duffy binding like gene; Duffy antigen binding protein; erythrocyte;
 KM DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;
 KM Plasmodium.
 XX
 OS Plasmodium vivax.
 OS Plasmodium falciparum.
 XX
 PN W09640766-A2.
 XX
 PD 19-DEC-1996.
 XX
 PF 07-JUN-1996; 96WO-US09508.
 XX
 PR 07-JUN-1995; 95US-0487826.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;
 XX
 DR WPI: 1997-052231/05.
 DR N-PSDB; AAT72882.

PT New malaria vaccines - contains cysteine-rich DBL family protein
 PT binding domains homologous domains of the Duffy and stialic acid
 PT binding proteins
 XX
 PS Claim 8; Page 61-67; 96pp; English.

CC This sequence represents var-7 of Plasmodium. Var-7 belongs to
 CC the Duffy binding like (DBL) family of genes which have homology to the
 CC Duffy antigen binding protein (DABP) and stialic acid binding protein
 CC (SABP) conserved regions (see AAY72889 and AAT72888 respectively). The
 CC var family of genes modulate cytoadherence and antigenic variation of
 CC Plasmodium infected erythrocytes. SABP and the Duffy antigen binding
 CC protein (DABP) are soluble proteins that appear in the culture
 CC supernatant after infected erythrocytes release merozoites. DABP and
 CC SABP mediate the binding of merozoites and schizonts to the erythrocyte
 CC surface. These proteins are necessary for erythrocyte invasion by the
 CC parasite. This sequence can be used in the compositions of the
 CC invention. The compositions are for the treatment and prevention of
 CC malaria, and comprise either a nucleotide sequence or encoded polypeptide
 CC of the var-1, var-2, var-3 or var-7 genes of the DBL gene family, a
 CC family of genes having homology with conserved regions of DABP and SABP.
 CC The compositions are used for the treatment and prevention of malaria.
 CC They are also used in the preparation of vaccines or induction of a
 CC protective immune response in a mammal to Plasmodium merozoites
 CC (especially Plasmodium falciparum or Plasmodium vivax).
 CC
 SO Sequence 3060 AA;

Query Match 39.2%; Score 669.5; DB 18; Length 3060;
 Best Local Similarity 40.9%; Pred. No. 1.1e-50;
 Matches 133; Conservative 54; Mismatches 109; Indels 29; Gaps 9;

QY 1 GNDGSENEISGNCNPKESYDPMDCKKNIDNSHGACMPRRKLC---VVDLTGGGELRKPE 57
 DB 1356 GKNQRTVGVGCNPKESYDPMDCKKNIDNSHGACMPRRKLC---VRLTGGGELRKPE 57

```

QY 58 DILTRFNCIAKETHFAWHKRYKKNVNAENELKSGKIPEGRKOMYTFGDFRDIFFCTD 117
   :: 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 1416 NLKDAFIKTAAEFLSMQYKSKNDSEAKILDRGLIPSOFLRSMMTTFGDIRDCLNTD 1475
QY 118 ISSCRYIKDTSQITKSKLG-----DQATTEKGDTHIDNKKLOEMWTIHGPKIEGMICA 172
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 1476 ISKQ--NDVAKA-KDKIGKFFSDGSKSPSGLSR-----QEMWKTNGPEIKGMICA 1525
QY 173 LTNLGISEE-KKNLIDYSYKKNLNAEKDCCLEKFAKPOFLRMVYEMSPFCRERKLL 231
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 1526 LTKVYVTDNKRRIKNDSTDKVNOSONGNSLEEFAPKPOFLRMVTEMGEFCAEROKK 1585
QY 232 EDKVEDVCIRAKDEYEGCKNNKNSCVKCEYENYITGKKTQYESQEGKF--NTEKRQK 289
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 1586 ENIIKDACNEINSTQOC--NDAKHRCNQACRAQOEYENKKEKFEFGGOTNNFVLKANVOPQ 1643
QY 290 KPEVNSYSKRD-----ASEYLKDK 308
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 1644 DPEYKGYEYKDGVOPIQGNEXYLQK 1668

```

RESULT 7

```

AA77905 ID AAT77905 standard; protein; 3060 AA.
AA77905 AC AAT77905;
DT 13-JUN-2000 (first entry)
XX DE Plasmodium var-7 polypeptide.
XX DE DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
KW DABP; Sialic Acid Binding Protein; SAMP; malaria; vaccine; immunisation;
KW protozoacide; var-7.
XX OS Plasmodium sp.
XX PN US5993827-A.
XX PD 30-NOV-1999.
XX PF 07-JUN-1995; 95US-0487826.
XX PR 10-SEP-1993; 93US-0119677.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

```

Sim KL, Chltnis C, Peterson DS, Su X, Wellens TE, Miller LH;

DR WPI: 2000-194198/17.
DR N-PSDB: AA298287.

PT Isolated protein binding domains from Plasmodium vivax and Plasmodium
PT falciparum erythrocyte binding proteins useful for vaccinating against
PT malaria -

PS Disclosure: Columns 109-124; 93pp; English.

XX The invention relates to ebl-1 polypeptides that are encoded by the DBL
CC (Duffy-binding like) gene family. The ebl-1 proteins are substantially
CC identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid
CC Binding Protein (SABP), which are soluble proteins that appear in the
CC culture supernatant after erythrocytes infected with malaria release
CC merozoites. Immunochemical studies indicate that DABP and SABP are the
CC respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy
CC and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be
CC used to vaccinate against malaria, especially caused by P. falciparum.
CC Immunisation with the polypeptide provides effective protection against
CC malaria. The present sequence represents the var-7 polypeptide.

SQ Sequence 3060 AA;

Query Match 39.2%; Score 669.5; DB 21; Length 3060;
Best Local Similarity 40.9%; Pred. No. 1.1e-50;
Matches 133; Conservative 54; Mismatches 109; Indels 29; Gaps 9;

```

QY 1 GNDGSNETSGCNPRESTPDMCKKINDNSHGACMPRRQKLC---VADLTQSGEIRKPE 57
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 1356 GKNRTYVGECPKESYPDMDCKNNIDISHGACMPRRQKLCYYIAHESQETENIKRD 1415
QY 58 DILTRFNCIAKETHFAWHKRYKKNVNAENELKSGKIPEGRKOMYTFGDFRDIFFCTD 117
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 1416 NLKDAFIKTAAEFLSMQYKSKNDSEAKILDRGLIPSOFLRSMMTTFGDIRDCLNTD 1475
QY 118 ISSCRYIKDTSQITKSKLG-----DQATTEKGDTHIDNKKLOEMWTIHGPKIEGMICA 172
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 1476 ISKQ--NDVAKA-KDKIGKFFSDGSKSPSGLSR-----QEMWKTNGPEIKGMICA 1525
QY 173 LTNLGISEE-KKNLIDYSYKKNLNAEKDCCLEKFAKPOFLRMVYEMSPFCRERKLL 231
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 1526 LTKVYVTDNKRRIKNDSTDKVNOSONGNSLEEFAPKPOFLRMVTEMGEFCAEROKK 1585
QY 232 EDKVEDVCIRAKDEYEGCKNNKNSCVKCEYENYITGKKTQYESQEGKF--NTEKRQK 289
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 1586 ENIIKDACNEINSTQOC--NDAKHRCNQACRAQOEYENKKEKFEFGGOTNNFVLKANVOPQ 1643
QY 290 KPEVNSYSKRD-----ASEYLKDK 308
   11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 1644 DPEYKGYEYKDGVOPIQGNEXYLQK 1668

```

RESULT 8

AAB62148 ID AAB62148 standard; peptide; 407 AA.

AC AAB62148;
DT 29-MAY-2001 (first entry)

XX P. falciparum varCSA polypeptide A4 DBL4-gamma.

KW FCR3.varCSA protein; chondroitin sulfate A; CSA; var gene; PFEPM1;
KW erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
KW malaria; protozoacide; A4 DBL4-gamma.

XX Plasmodium falciparum.

XX WO200116326-A2.

XX 08-MAR-2001.

XX 01-SEP-2000; 2000WO-US24195.

XX 01-SEP-1999; 99US-0152023.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;
PI Gysin J, Pouvelle B, Fujii N, Smith J;

DR WPI: 2001-235109/24.

PT Novel FCR3.varCSA protein, useful for modulating parasitized red blood
PT cell binding, sequestration and onset of maternal malaria -

PS Disclosure: Page 72-73; 78pp; English.

XX The invention relates to a P. falciparum FCR3.varCSA protein, that is
CC capable of binding to chondroitin sulfate A (CSA). The var gene and the
CC corresponding P.falciparum erythrocyte membrane protein 1 (PFEPM1)
CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
CC protein and the encoding gene are useful for treating and preventing
CC maternal malaria in a patient identified at a risk for contracting
CC maternal malaria or in a patient afflicted with maternal malaria. The
CC present sequence represents a P. falciparum varCSA polypeptide

[illegible]

XX Erythrocyte binding ligand (EBL) family genes were cloned from
 CC P. falciparum chromosome 7 subsegment libraries constructed during
 CC genetic studies of the chloroquine resistance locus. The 4 genes,
 CC EBL-e1 (AA083526), E3ia (AA083527), EBL-e2 (AA083528) and Pro3
 CC (AA083529), encode the proteins given in AAT70233-36, respectively. The
 CC binding domains of such proteins can be expressed e.g. in E. coli,
 CC yeast, mammalian, insect, and in vaccinia virus and adenovirus-infected
 CC cells, and provide protection against P. falciparum.
 XX

SO Sequence 793 AA:
 Query Match 30.5%; Score 521; DB 16; Length 793;
 Best Local Similarity 37.9%; Pred. No. 4.2e-38;
 Matches 127; Conservative 46; Mismatches 106; Indels 56; Gaps 13;

QY 3 DGSNETSGCNPKESYDPMC-KKNIDNSHGACMPRRKQCLVRDLTGGGEIRKPPDILT 61
 DB 407 NGRTTVGECYRKETYSMTCDSEKIMGOGACIPRRKQCLVH------EKIMT 457
 DB 62 K-----FINCAKETHFAMHKYKRD-NVNAEN---ELKSGKIPGFRKMYTTEGDR 110
 DB 458 NTNELKATFIKCAAAETFLMONTKKDKNGNAEDDEKLKGLTFPDRKMTYTDYR 517
 QY 111 DIFFGTDISCRYIKDTSQTIKSLGDAQTEKGDTHID-----NKKLQEWMTIHG 162
 DB 518 DICLGTDIS--KKDTSKGV-----GKVCNIDDVFKYSINSIRKSMWETNG 564
 QY 163 PRTWEGMLCALTNGLS-----ESEKKNILQDYSYKLNNAEKDDCCLEKFAKPPFLRW 216
 DB 565 PRTWEGMLCALSTYDTSLNVPETHRKLTEGNNNEKVIYFGSDSFTLSKFSERPOFLRW 624
 QY 217 YVWSEDFCERERKRLDKVEDVCIRAKDEYEGCKNNKSNNSCVV--CKEYENYITGKRT 273
 DB 625 LTFWGMFPCFEQCK-EYKV--LLAKCKDDVDGDGCKNGKVCACRQCKQYHSHWIGIWD 681
 QY 274 QYESQEGKFNTEKR--QKKPEYNSYSKKDASEYLK 306
 DB 682 NYKKQGRYTEVKKIPLYKEDKDVKNSDADRYLK 716

RESULT 11
 AAM22480
 ID AAM22480 standard; Protein; 921 AA.
 XX AC AAM22480;
 XX 07-OCT-1997 (first entry)
 XX Plasmodium E3ia.
 XX DBL gene family; SABP; sialic acid binding protein; vaccine; therapy;
 KM Duffy binding like gene; Duffy antigen binding protein; erythrocyte;
 KM DABP; merozoite; malaria; var-1; var-2; var-3; var-7; Immune response;
 KM Plasmodium.
 XX Plasmodium falciparum.
 OS MO9640766-A2.
 XX PN 19-DEC-1996.
 XX PD 07-JUN-1996; 96WO-US09508.
 XX PF 07-JUN-1996; 96WO-US09508.
 XX PR 07-JUN-1995; 95US-0487826.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;
 PI Wellens TE;
 XX WPI: 1997-052231/05.
 DR N-PSDB: AAT72895.

XX New malaria vaccines - contains cysteine-rich DBL family protein
 PT binding domains homologous domains of the Duffy and sialic acid
 PT binding proteins
 XX Disclosure; Page 43-45; 96pp; English.
 XX This sequence represents E3ia of Plasmodium. E3ia belongs to
 CC the Duffy binding like (DBL) family of genes which have homology to the
 CC Duffy antigen binding protein (DABP) and sialic acid binding protein
 CC (SABP) conserved regions (see AAT72889 and AAT72888 respectively). The
 CC var family of genes modulate cytoadherence and antigenic variation of
 CC Plasmodium infected erythrocytes. SABP and the Duffy antigen binding
 CC protein (DABP) are soluble proteins that appear in the culture
 CC supernatant after infected erythrocytes release merozoites. DABP and
 CC SABP mediate the binding of merozoites and schizonts to the erythrocyte
 CC surface. These proteins are necessary for erythrocyte invasion by the
 CC parasite. This sequence can be used in the compositions of the
 CC invention. The compositions are for the treatment and prevention of
 CC malaria, and comprise either a nucleotide sequence or encoded polypeptide
 CC of the var-1, var-2, var-3 or var-7 genes of the DBL gene family.
 CC family of genes having homology with conserved regions of DABP and SABP.
 CC The compositions are used for the treatment and prevention of malaria.
 CC They are also used in the preparation of vaccines for inducing a
 CC protective immune response in a mammal to Plasmodium merozoites
 CC (especially Plasmodium falciparum or Plasmodium vivax).
 XX

SO Sequence 921 AA:
 Query Match 30.5%; Score 521; DB 18; Length 921;
 Best Local Similarity 37.9%; Pred. No. 5.2e-38;
 Matches 127; Conservative 46; Mismatches 106; Indels 56; Gaps 13;

QY 3 DGSNETSGCNPKESYDPMC-KKNIDNSHGACMPRRKQCLVRDLTGGGEIRKPPDILT 61
 DB 407 NGRTTVGECYRKETYSMTCDSEKIMGOGACIPRRKQCLVH------EKIMT 457
 DB 62 K-----FINCAKETHFAMHKYKRD-NVNAEN---ELKSGKIPGFRKMYTTEGDR 110
 DB 458 NTNELKATFIKCAAAETFLMONTKKDKNGNAEDDEKLKGLTFPDRKMTYTDYR 517
 QY 111 DIFFGTDISCRYIKDTSQTIKSLGDAQTEKGDTHID-----NKKLQEWMTIHG 162
 DB 518 DICLGTDIS--KKDTSKGV-----GKVCNIDDVFKYSINSIRKSMWETNG 564
 QY 163 PRTWEGMLCALTNGLS-----ESEKKNILQDYSYKLNNAEKDDCCLEKFAKPPFLRW 216
 DB 565 PRTWEGMLCALSTYDTSLNVPETHRKLTEGNNNEKVIYFGSDSFTLSKFSERPOFLRW 624
 QY 217 YVWSEDFCERERKRLDKVEDVCIRAKDEYEGCKNNKSNNSCVV--CKEYENYITGKRT 273
 DB 625 LTFWGMFPCFEQCK-EYKV--LLAKCKDDVDGDGCKNGKVCACRQCKQYHSHWIGIWD 681
 QY 274 QYESQEGKFNTEKR--QKKPEYNSYSKKDASEYLK 306
 DB 682 NYKKQGRYTEVKKIPLYKEDKDVKNSDADRYLK 716

RESULT 12
 AAT77902
 ID AAT77902 standard; Protein; 921 AA.
 XX AC AAT77902;
 XX 13-JUN-2000 (first entry)
 XX P. falciparum ebl-1 related polypeptide.
 XX DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
 KM DABP; Sialic Acid Binding Protein; SABP; malaria; vaccine; immunisation;
 KM protozoacide.
 XX Plasmodium falciparum.
 OS

XX US593827-A.
 PN
 XX
 PD 30-NOV-1999.
 XX
 XX 07-JUN-1995; 9505-0487826.
 PF
 XX 10-SEP-1993; 9305-0119677.
 PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX
 PI Sim KL, Chitnis C, Peterson DS, Su X, Wellens TE, Miller LH;
 XX
 XX WPI; 2000-194198/17.
 DR N-PSDB; AA698284.
 XX
 PT Isolated protein binding domains from Plasmodium vivax and Plasmodium
 PT falciparum erythrocyte binding proteins useful for vaccinating against
 PT malaria -
 Disclosure: Columns 61-66; 93pp; English.

CC The invention relates to ebl-1 polypeptides that are encoded by the DBL
 CC (Duffy-binding like) gene family. The ebl-1 proteins are substantially
 CC identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid
 CC Binding Protein (SABP), which are soluble proteins that appear in the
 CC culture supernatant after erythrocytes infected with malaria release
 CC merozoites. Immunochemical studies indicate that DABP and SABP are the
 CC respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy
 CC and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be
 CC used to vaccinate against malaria, especially caused by P. falciparum.
 CC Immunization with the polypeptide provides effective protection against
 CC malaria.
 CC
 XX
 SQ Sequence 921 AA;

Query Match 30.5%; Score 521; DB 21; Length 921;
 Best Local Similarity 37.9%; Pred. No. 5.2e-38;
 Matches 127; Conservative 46; Mismatches 106; Indels 56; Gaps 13;

QY 3 DGSNEISGCPNPKESYPMDC-KKNIDNSHSGACMPPRKOKCYVDLGGGIRKPEDILT 61
 DB 407 NGRTTGVGCRKREYSEMTODESKIKKGHCACIPPRKOKICLHYL-----EKIMT 457
 QY 62 K-----FINCAKETHPFAMHKYKQD-NVNAEN--ELSGKLPFGFRKQMYTFEGDFR 110
 DB 458 NFNELKYAFIKCAAEFTFLMQNYKKDKNGNAEDLDEKLGKGIPEDFKQMFYTFADYR 517
 QY 111 DIFEGTDISSCRVYKTSQITKSLGDAQTEKGDTHID-----NKKLOEMWTIHG 162
 DB 518 DICLGTDISS--KKDTSKGV-----GKVCNIDIVFYKISINSIRYKRSKWETNG 564
 QY 163 PRTWEGMICALTNGLS-----ESEKNILQDYSYNKLNNAEKDDCLFEFASKPOFLRW 216
 DB 565 PVTWEGMICALSYDTSINNVNPEHTKKLTGEGNNNEFEVIRGSSSTLSKFSERPOFLRW 624
 QY 217 YVWSDFECEERKRLKLEKVEDVCIAKADYGCNNKNSNNSCVV--CKEYENYITGKKT 273
 DB 625 LTEWGENFCEQKK-EKKV--LLAKCKDQVDGDGKNGKCVACKDQCKQYHSMIGIWD 681
 QY 274 QYESQEGFNTERR--OKRPEVNSYSKKDASEYLK 306
 DB 682 NYKKQKRGTYEVKAKIPIYKEDKDKVNSDARDYTK 716

RESULT 13
 AAB62147
 ID AAB62147 standard; peptide; 294 AA.
 XX
 AC AAB62147;
 XX
 DT 29-MAY-2001 (first entry)
 XX

DE P. falciparum varCSA polypeptide R29DBL2-gamma.

XX FCR3, varCSA protein; chondroitin sulfate A; CSA; var gene; PFEMP1;
 KW erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
 KW malaria; protozoacide; R29DBL2-gamma.
 XX

OS Plasmodium falciparum.

PN WO200116326-A2.

PD 08-MAR-2001.

PE 01-SEP-2000; 2000WO-US24195.

PR 01-SEP-1999; 99US-0152023.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;
 PI Gysin J, Pouvelle B, Fujii N, Smith J;
 XX
 XX WPI; 2001-235109/24.

DR Novel FCR3, varCSA protein, useful for modulating parasitized red blood
 DR cell binding, sequestration and onset of maternal malaria -
 XX
 XX Disclosure: Page 72; 78pp; English.

CC The invention relates to a P. falciparum FCR3, varCSA protein, that is
 CC capable of binding to chondroitin sulfate A (CSA). The var gene and the
 CC corresponding P. falciparum erythrocyte membrane protein 1 (PEMP1)
 CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
 CC protein and the encoding gene are useful for treating and preventing
 CC maternal malaria in a patient afflicted at a risk for contracting
 CC maternal malaria or in a patient afflicted with maternal malaria. The
 CC present sequence represents a P. falciparum varCSA polypeptide
 CC R29DBL2-gamma.
 CC
 XX
 SQ Sequence 294 AA;

Query Match 30.1%; Score 514; DB 22; Length 294;
 Best Local Similarity 37.4%; Pred. No. 4.7e-38;
 Matches 110; Conservative 48; Mismatches 112; Indels 24; Gaps 11;

QY 5 SNEISGCPNPKESYPMDC-KKNIDNSHSGACMPPRKOKCYVDLGGGIRKPEDILT 61
 DB 7 NSIDNCAKNNKRRKNEMOCCKNTFVNGEGVCMPPRRKSCICHNLLEQOTKNTYQLREA 66
 QY 63 FINCAKETHPFAMHKYKQD-NVNAEN--ELSGKLPFGFRKQMYTFEGDFRIFGTDISSCR 122
 DB 67 FTICAKKETNLLMDKDKKNDKNEAEILKKGKTIPEDFMRIMFTFDFRDCLENDMG--- 123
 QY 123 YIKDTSQITKSLGD--QATTEKGTHTIDNNKLOEMWTIHGPKTWEGMICALTNGLS 180
 DB 124 --KVDYK-VKKNINKVFNNSKRGFKIDP---ENMWNENGPQIWMGICALHIAHDKD 176
 QY 181 EKKNIQDYSYNKLNNAEKDD--CJLEKFAKPOFLRWYVWSDFECEERKRLKLEKVE 236
 DB 177 SLKN-KDNKYKRYVITLLAKRDGSGNMTLSEFKKRPFLRWFEWYDDDCREROKYLTTEVA 235
 QY 237 DVCIRAKDYEGCKNNKNSNNSCVKCKEYENYITGKTOYESQEGFNTERRKOK 290
 DB 236 STC---KSIDG-GQLKCDRGCKNNKCKDEKRYKRRKKKEEMNQD-KYYDKRENK 284

RESULT 14
 AAB62149
 ID AAB62149 standard; peptide; 351 AA.
 XX
 AC AAB62149;
 XX
 DT 29-MAY-2001 (first entry)
 XX

DE P. falciparum varCSA polypeptide A4tresDBL3-gamma.
 XX
 XX FCR3, varCSA protein; chondroitin sulfate A; CSA; var gene; PFEMP1;
 KW erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
 XX malaria; protozoacide; A4tresDBL3-gamma.
 OS Plasmodium falciparum.
 XX
 PN W0200116326-A2.
 XX
 PD 08-MAR-2001.
 XX
 PF 01-SEP-2000; 2000MO-US24195.
 XX
 PR 01-SEP-1999; 99US-0152023.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Scherf A, Miller LH, Gamain B, Barnuch DI, Buffet P, Scheidig C;
 DR Gysin J, Pouvelle B, Fujii N, Smith J;
 WPI; 2001-235109/24.
 XX
 XX Novel FCR3, varCSA protein, useful for modulating parasitized red blood
 PT cell binding, sequestration and onset of maternal malaria -
 XX
 PS Claim 54; Page 73-74; 78pp; English.
 XX
 CC The invention relates to a P. falciparum FCR3, varCSA protein, that is
 CC capable of binding to chondroitin sulfate A (CSA). The var gene and the
 CC corresponding P.falciparum erythrocyte membrane protein 1 (PFEMP1)
 CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
 CC protein and the encoding gene are useful for treating and preventing
 CC maternal malaria in a patient afflicted at a risk for contracting
 CC maternal malaria or in a patient afflicted with maternal malaria. The
 CC present sequence represents a P. falciparum varCSA polypeptide
 CC A4tresDBL3-gamma.
 CC
 XX
 XX Sequence 351 AA:
 SO
 Query Match 26.08; Score 444; DB 22; Length 351;
 Best Local Similarity 34.48; Pred. No. 1.2e-31;
 Matches 115; Conservative 46; Mismatches 131; Indels 42; Gaps 12;
 YY 6 NEISGNGPKREYRDWCKKIDNSHGACMPRRROKLCVADLQSGEIRKPELILKFIN 65
 DB 16 NNHNCKKTED AKMCCENTKLGEDRGVCMPPRRQNLVHYLRKLNDSKEEDLREAFIK 74
 66 CAKEEYHFAWHKRYKKDNVNAENELKSGKIPEGRKQMYTFEGDFRDIFFGTDISCRYIK 125
 75 SAARFELRLQYNSKNVEDDKILHRDMLPPEFRSMFYTFGRDCLDTDISSEKIAKH 134
 YY 126 DTSQITKSKIGDQATEKGDTHIDNKKLOE---WMTINGPKIWEGLCALJNGLSESEK 182
 DB 135 DVT-TAKKKT--TAVFOKIGSKTNGKYLERGWKEVGLSTIMKMLCALSYN-TETKK 190
 YY 183 -----NNIIDDYSYNLNAEKDDCCLEKFAKPOELRWVENSDECRERKK-----LED 233
 DB 191 MDECVRIYLMKTYIK--NNDIKF--YLFESASRPPLRWLTWEGEDVKNRKKELVSLK 246
 YY 234 KVEDVCIR--AKDYGCKNNKSNNSCVKCEYENYITGKKTQYESQGRF-----282
 DB 247 KCSQCTLRNNGTSNKTCDODNENCGACKTQCEKTKKMMERKKHYSQKKKFOLYKNSATY 306
 YY 283 -----NTERKQKKPEY-NSYSKKDASEYIK 306
 DB 307 NNGIAYKEANSEYTKNDPEVTEANSASAKHARDYIK 340

AC AAM93944;
 XX
 XX 30-JUN-1999 (first entry)
 DE P. falciparum PFEMP1 protein.
 XX
 KW Erythrocyte membrane protein; EMP; PFEMP1; malaria; antioocclusion;
 KW glycosamino-glycan-like moiety; antiaggregational; antimalarial;
 KW antigen receptor; infected erythrocyte; rosette formation; blood cell;
 KW capillary occlusion; cerebral malaria; treatment; vaccine; detection;
 KW medicament; parasite; diagnosis; drug screening.
 OS Plasmodium falciparum.
 XX
 PN W09915557-A1.
 XX
 PD 01-APR-1999.
 XX
 PF 18-SEP-1998; 98MO-SE01675.
 XX
 PR 19-SEP-1997; 97SE-0003386.
 XX
 PA (KARO-) KAROLINSKA INNOVATIONS AB.
 XX
 PI Barragan A, Carlson J, Fernandez V, Qijun C, Wahlgren M;
 DR WPI; 1999-254692/21.
 XX
 XX New isolated malaria polypeptides.
 PT
 PS Claim 4; Page 67-74; 80pp; English.
 XX
 CC This invention describes a novel Plasmodium falciparum erythrocyte
 CC membrane protein (EMP), PFEMP1, which is capable of binding to a
 CC carbohydrate which exhibits at least one negatively charged
 CC glycosamino-glycan (GAG)-like moiety and has antiaggregational,
 CC antioocclusion and antimalarial activity. The carbohydrates of the
 CC invention are capable of acting as receptors for malaria antigens
 CC present on the surfaces of malaria infected erythrocytes, by binding
 CC to these antigens the carbohydrates prevent rosette formation by the
 CC blood cells, this prevents occlusion of capillaries as is seen in
 CC cerebral malaria caused by Plasmodium falciparum. The products of the
 CC invention can be used to treat malaria or to vaccinate against it, or
 CC used to design a model to identify compounds that bind to PFEMP1. The
 CC carbohydrates, polypeptides and antibodies of the invention can be used
 CC as a medicament for dissolving the rosettes formed by erythrocytes
 CC infected by a malaria parasite. The products can also be used for
 CC detection, diagnosis and drug screening.
 CC
 XX
 XX Sequence 2228 AA:
 SO
 Query Match 22.68; Score 386; DB 20; Length 2228;
 Best Local Similarity 24.58; Pred. No. 2.6e-25;
 Matches 104; Conservative 59; Mismatches 123; Indels 138; Gaps 14;
 YY 20 WDC-----KKINDNSHGA-CMPRRROKLCVRL-----TGGGIRK 55
 DB 917 WRCVTTSGEPTTSSDKKGAICVPPRRRLTIKTIQVMATESPQASGSASTSGSTTP 976
 YY 56 P---EDILTFKINCAKETHFAWHKRYKKD-----NVNAENELKS 91
 DB 977 PDSKELALKAFVESALETFFLWHRYKEKKAVAOEGAGHGLRVERGSSPEYDEPKLKE 1036
 YY 92 GKTIPEGRKQMYTFEGDFRDIFFGTDISCRYIKDTSQITKSKIGDQATEKGDTHIDN 151
 DB 1037 GKIPDGLRLQMYTFGLDGYRDLFGSNDTTSVSKDTPSSSDNLKNTVLLASGSTE-QER 1095
 YY 152 KRL-----QEWMTIHGPKIWEGLCALF-----174
 DB 1096 EKMKYKEIKNFKCSERSAPNLVSHPQTWENRNGKTYIWHGNVCAITLSKDTAKGYEKK 1155
 YY 175 -----NGLSEKKNNIIDDYSYNKLN-----NAEKDDCLEKFAK 210

RESULT 15
 AAM93944
 ID AAM93944 standard; protein; 2228 AA.

| | | | |
|----|------|---|-------|
| Db | 1156 | PQKLENNENLMDANKKPKPPQYQYTNKKLDENSGTSPRTQYQASSNPTTTLTHPVKR | 12119 |
| QY | 211 | POFLRWYEWSDFFCRRKK-----KLEDAVE--DYCIRAKADYEGCKNKNKSNN----- | 255 |
| Db | 1216 | PTYRRFEWEEESFCRRKKRLKQIKVDCKEENVDGCSGDGCADCSISTHDYSTVSEF | 1275 |
| QY | 256 | ---SCVAVCKRLEYNTJGKTKQYESQCEKFNTEKRQ--KKPEYNSYSKK-----DAS | 302 |
| Db | 1276 | NCPCGCKGRKSSRYKWIERRKTEFHKQSNVAYGQOKTDAUVRNNGNTFDEKCFKLTETWPDAA | 1335 |
| QY | 303 | EYK 306 | |
| Db | 1336 | KFL 1339 | |

Search completed: June 20, 2003, 15:02:56
Job time : 50.5417 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 15:01:19 ; Search time 16.8255 Seconds
(without alignments)
538.603 Million cell updates/sec

Title: US-10-087-013-11

Perfect score: 1706
Sequence: 1 GNDGSENEISGNCNPKESYPDM.....KKPEVNSYSKKDASEYLKDK 308

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-------------------|--------------------|
| 1 | 669.5 | 39.2 | 2710 | 2 | US-08-568-459A-12 | Sequence 12, Appl |
| 2 | 669.5 | 39.2 | 2710 | 2 | US-08-487-826B-12 | Sequence 12, Appl |
| 3 | 669.5 | 39.2 | 2710 | 4 | US-09-210-288-12 | Sequence 12, Appl |
| 4 | 669.5 | 39.2 | 3060 | 2 | US-08-487-826B-14 | Sequence 14, Appl |
| 5 | 521 | 30.5 | 921 | 2 | US-08-568-459A-8 | Sequence 8, Appl |
| 6 | 521 | 30.5 | 921 | 2 | US-08-487-826B-8 | Sequence 8, Appl |
| 7 | 521 | 30.5 | 921 | 4 | US-09-210-288-8 | Sequence 8, Appl |
| 8 | 367.5 | 21.5 | 2182 | 2 | US-08-487-826B-16 | Sequence 16, Appl |
| 9 | 268.5 | 15.7 | 749 | 2 | US-08-568-459A-6 | Sequence 6, Appl |
| 10 | 268.5 | 15.7 | 749 | 2 | US-08-487-826B-6 | Sequence 6, Appl |
| 11 | 268.5 | 15.7 | 749 | 4 | US-09-210-288-6 | Sequence 6, Appl |
| 12 | 263.5 | 15.4 | 311 | 2 | US-08-568-459A-21 | Sequence 21, Appl |
| 13 | 263.5 | 15.4 | 311 | 2 | US-08-487-826B-33 | Sequence 33, Appl |
| 14 | 263.5 | 15.4 | 311 | 4 | US-09-210-288-21 | Sequence 21, Appl |
| 15 | 259 | 15.2 | 1115 | 2 | US-08-568-459A-2 | Sequence 2, Appl |
| 16 | 259 | 15.2 | 1115 | 2 | US-08-487-826B-2 | Sequence 2, Appl |
| 17 | 259 | 15.2 | 1115 | 4 | US-09-210-288-2 | Sequence 2, Appl |
| 18 | 259 | 15.2 | 1115 | 6 | 5198347-6 | Patent No. 5198347 |
| 19 | 251 | 14.7 | 1435 | 2 | US-08-568-459A-4 | Sequence 4, Appl |
| 20 | 251 | 14.7 | 1435 | 2 | US-08-487-826B-4 | Sequence 4, Appl |
| 21 | 251 | 14.7 | 1435 | 4 | US-09-210-288-4 | Sequence 4, Appl |
| 22 | 237.5 | 13.9 | 700 | 2 | US-08-568-459A-10 | Sequence 10, Appl |
| 23 | 237.5 | 13.9 | 700 | 2 | US-08-487-826B-10 | Sequence 10, Appl |
| 24 | 237.5 | 13.9 | 700 | 4 | US-09-210-288-10 | Sequence 10, Appl |
| 25 | 221 | 13.0 | 324 | 2 | US-08-568-459A-17 | Sequence 17, Appl |
| 26 | 221 | 13.0 | 324 | 2 | US-08-487-826B-29 | Sequence 29, Appl |
| 27 | 221 | 13.0 | 324 | 4 | US-09-210-288-17 | Sequence 17, Appl |

| | | | | | | |
|----|-------|------|------|---|-------------------|--------------------|
| 28 | 178 | 10.4 | 197 | 6 | 5198347-2 | Patent No. 5198347 |
| 29 | 178 | 10.4 | 778 | 6 | 5198347-4 | Patent No. 5198347 |
| 30 | 156 | 9.1 | 277 | 2 | US-08-568-459A-15 | Sequence 15, Appl |
| 31 | 156 | 9.1 | 277 | 2 | US-08-487-826B-27 | Sequence 27, Appl |
| 32 | 136 | 9.1 | 277 | 4 | US-09-210-288-15 | Sequence 15, Appl |
| 33 | 135 | 7.9 | 291 | 2 | US-08-568-459A-13 | Sequence 13, Appl |
| 34 | 135 | 7.9 | 291 | 2 | US-08-487-826B-25 | Sequence 25, Appl |
| 35 | 135 | 7.9 | 291 | 4 | US-09-210-288-13 | Sequence 13, Appl |
| 36 | 127.5 | 7.5 | 282 | 2 | US-08-568-459A-16 | Sequence 16, Appl |
| 37 | 127.5 | 7.5 | 282 | 2 | US-08-487-826B-28 | Sequence 28, Appl |
| 38 | 127.5 | 7.5 | 282 | 4 | US-09-210-288-16 | Sequence 16, Appl |
| 39 | 119.5 | 7.0 | 362 | 2 | US-08-568-459A-18 | Sequence 18, Appl |
| 40 | 119.5 | 7.0 | 362 | 2 | US-08-487-826B-30 | Sequence 30, Appl |
| 41 | 119.5 | 7.0 | 362 | 4 | US-09-210-288-18 | Sequence 18, Appl |
| 42 | 112 | 6.6 | 411 | 2 | US-08-568-459A-19 | Sequence 19, Appl |
| 43 | 112 | 6.6 | 411 | 2 | US-08-487-826B-31 | Sequence 31, Appl |
| 44 | 112 | 6.6 | 411 | 4 | US-09-210-288-19 | Sequence 19, Appl |
| 45 | 109 | 6.4 | 2391 | 2 | US-08-446-855A-2 | Sequence 2, Appl |

ALIGNMENTS

RESULT 1
US-08-568-459A-12
; Sequence 12, Application US/08568459A
; Patent No. 5843306
; GENERAL INFORMATION:
; APPLICANT: Slim, Kim L.
; APPLICANT: Chintis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhuan
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelson, Ned
; REGISTRATION NUMBER: 29,655
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-0176
; TELEFAX: (619) 235-8550
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2710 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; US-08-568-459A-12
Query Match 39.2%; Score 669.5; DB 2; Length 2710;

Best Local Similarity 40.9%; Pred. No. 6.7e-53;
Matches 133; Conservative 54; Mismatches 109; Indels 29; Gaps 9;

```
OY 1 GNDGSENEISGCPKESYPDMCKNNIDSHSGACMPPRKRLC---VADLTGGGEIRKPE 57
Db 1358 GRNGTTTGVGECNPKESYPDMCKNNIDSHSGACMPPRKRLCYYIAHESQTENIKTD 1417
OY 58 DILTKFNCARKEHFAMHKKYKKNVNAENELSGKPIEGEPRKQMYTFEGDFDIFGCTD 117
Db 1418 NLKDAFRTAAAEFLSMQYKSKNDSEAKILDRGLIPSOFLRSMMTFFGIDYRDLCLNTD 1477
OY 118 ISSCRYIKDTSQTIKSLG-----DQATTEKGDTHIDNKKLQEWMTIHGPKIWEGLCA 172
Db 1478 ISKKQ--NDVAKA-KDKIGKFFSKDGSKSPSGLSR-----QEWMTINGPETIKWGMCA 1527
OY 173 LTNLSESE-KKNILQDYSTNKLNNAEKDDCCLEKFAKPOFLRMVYVMSDFCERERKTL 231
Db 1528 LTKVYTDPNKRRIKNDYSYDKVNOSONGNSLEEFAPKPOFLRMVMEWGEFCARERKK 1587
OY 232 EDKVEDYCIKAKDYEGCKNNKSNVCYKVEYENYITGKKTQYESQEGKF--NTEKRQK 289
Db 1588 ENIKKACNEINSTQOC--NDAKHRCNQCRAVOEYENKKKEFSQGTNNFVLKANVQPO 1645
OY 290 KPEYNSYSKRD-----ASEYLRK 308
Db 1646 DPEYKGYEKDGVQPIQGNERYLLQK 1670
```

RESULT 2

US-08-487-826B-12
Sequence 12, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
APPLICANT: Chlunis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH21.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO

ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-08-487-826B-12

Query Match 39.2%; Score 669.5; DB 2; Length 2710;
Best Local Similarity 40.9%; Pred. No. 6.7e-53;
Matches 133; Conservative 54; Mismatches 109; Indels 29; Gaps 9;

```
OY 1 GNDGSENEISGCPKESYPDMCKNNIDSHSGACMPPRKRLC---VADLTGGGEIRKPE 57
Db 1358 GRNGTTTGVGECNPKESYPDMCKNNIDSHSGACMPPRKRLCYYIAHESQTENIKTD 1417
OY 58 DILTKFNCARKEHFAMHKKYKKNVNAENELSGKPIEGEPRKQMYTFEGDFDIFGCTD 117
Db 1418 NLKDAFRTAAAEFLSMQYKSKNDSEAKILDRGLIPSOFLRSMMTFFGIDYRDLCLNTD 1477
OY 118 ISSCRYIKDTSQTIKSLG-----DQATTEKGDTHIDNKKLQEWMTIHGPKIWEGLCA 172
Db 1478 ISKKQ--NDVAKA-KDKIGKFFSKDGSKSPSGLSR-----QEWMTINGPETIKWGMCA 1527
OY 173 LTNLSESE-KKNILQDYSTNKLNNAEKDDCCLEKFAKPOFLRMVYVMSDFCERERKTL 231
Db 1528 LTKVYTDPNKRRIKNDYSYDKVNOSONGNSLEEFAPKPOFLRMVMEWGEFCARERKK 1587
OY 232 EDKVEDYCIKAKDYEGCKNNKSNVCYKVEYENYITGKKTQYESQEGKF--NTEKRQK 289
Db 1588 ENIKKACNEINSTQOC--NDAKHRCNQCRAVOEYENKKKEFSQGTNNFVLKANVQPO 1645
OY 290 KPEYNSYSKRD-----ASEYLRK 308
Db 1646 DPEYKGYEKDGVQPIQGNERYLLQK 1670
```

RESULT 3

US-09-210-286-12
Sequence 12, Application US/09210286
Patent No. 6392026

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
APPLICANT: Chlunis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,286
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH21.1FWDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FWV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 921 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-09-210-288-8

Query Match 30.5%; Score 521; DB 4; Length 921;
Best Local Similarity 37.9%; Pred. No. 8.7e-40;
Matches 127; Conservative 46; Mismatches 106; Indels 56; Gaps 13;

QY 3 DGSNEISGNCNPKESYPDMDC--KKNIDNSHSGACMPPRROKLCVRLDLOGEIRRPEDILT 61
DB 407 NGRTTGVGCYKRETYSEWTCDESKIKMGQHGACIPRRQKLCIHL-----EKIMT 457
QY 62 K-----FICAKAEHFAMHKYKRD-NVAEN---ELKSGKIPEGRKQMYTFGDER 110
DB 458 NTNELKAFIKCAAEFFLLQWNYKKKNGNAEDLDEKLGIIPEDEKRMFTYFADYR 517
QY 111 DIFGTDISSCRITKDTISQTSKSLGQDATTEKGTHTID-----NKKLOEMWTIHG 162
DB 518 DICGTDISS--KKDTSKGV-----GKYKCNIDVDFYKINSIRTRKSMWETNG 564
QY 163 PKINEGMCALITNGLS-----ESEKNILQDIYSYNKLNNAEKDCCLEKFAKRPQLRW 216
DB 565 PVMEGMCALSYDTSLLNNVNPETHKKLLEGNNNEKVIFFGSDSTLTSKFERPQLRW 624
QY 217 YVENSDECRERKKLEDEKVEDVCIKADYEGCKNNKSNVCYK---CKEYENITGKKT 273
DB 625 LTEWGENFCBOKR-EYKV--LLAKCKDCDVGDKCKGCKVACCKDCKOYHSMGIWD 681
QY 274 QYESQEGKFNTER--QKPEYNSYSKKDASEYLK 306
DB 682 NYKKQKGRYTEVKRIPLYKEDKDYKNSDARDYIK 716

RESULT 8
US-08-487-826B-16
Sequence 16, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhuan
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach

STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2182 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-487-826B-16

Query Match 21.5%; Score 367.5; DB 2; Length 2182;
Best Local Similarity 25.4%; Pred. No. 4.7e-25;
Matches 107; Conservative 60; Mismatches 128; Indels 127; Gaps 16;

QY 10 GGNPKESYPDMDC--KKNIDNSHSG---ACMPPRROKLCVRLDLO---GGE----- 52
DB 874 GPGKEKPPNKKCVTPSGVSTATSGKDGALCVPPRRRLYVGGLSQWASRGDETFEYSS 933
QY 53 -----IRKPEDILTRKINCAKEHFAMHKYKRD-----NKKLOEMWTIHG 83
DB 934 EATNAPSQSESEKLTATIESAAIETFFLWHKYEKKRPATODGAGIGVSLPPSPPE 993
QY 84 NAENEL-KSGKIPGFRKQMYTFGDFRDIFF-GTDISSCRYIKDTISQTSKSLGQDATT 141
DB 994 DPOTLOQGTGVIIPDPLRQMTYTLADYKDLIYSSGNDTSDTGTOTBSSNDLNKITYLE 1093
QY 142 EKGDTHTIDNK-----KLOEMWTIHGPKIWEGMICA 172
DB 1054 ASGSTGEQEKEMKQIQAKIKILNGATSGVPPVTKNSVKTPOQTWMEIADINWAMYCA 1113
QY 173 LT-----NGIS-----ESEKNILQDIYSYNKLNNAEKDCC----- 202
DB 1114 LTYKENDARGTSANIEQNKDLKALWDEANKRPTIEKYQYNNVLEDESGAKSDTTOPP 1173
QY 203 CLKEFASKPQPLRWYVENSDECFERKKLEDEKVEDVCIK---AKDYEG---CKNNKS-- 253
DB 1234 YNVLDLSSSCAKRCILKTYIEKKKTEYEKQKQAYEQKSNYENEGDKCOTQSNNNAN 1293
QY 1174 TLKNFVELPTFRHLHENGNSFCFERAKRLAQIHCEKDEDEGEKQYSDGEYCEIFRSQ 1293
QY 254 -----NNSCVKCKEYENITGKKTQYESQEGKFNTER---RQKPEYNSYSKKDAS 302
DB 303 EY 304
DB 1294 EF 1295

RESULT 9
US-08-568-459A-6
Sequence 6, Application US/08568459A
Patent No. 5849306
GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VITAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568,459A
FILING DATE: 07-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 749 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-08-568-459A-6

Query Match 15.7%; Score 268.5; DB 2; Length 749;
Best Local Similarity 26.8%; Pred. No. 1.6e-16;
Matches 85; Conservative 45; Mismatches 116; Indels 71; Gaps 15;

6 NEISGCKNPKES--YPMDC-KKNDNSHGACMPPRROKLC-----VRDLTGGELRRP 56
109 SSVFGCKTKISKYKKKKKNCYSNNKYTKPEVCPPRROQLGTYIFLRDNGEG-----163
57 EDILTKFNCACAKETHFAWH-KYKKDNVNAENELKSGKIPGFRKOMYTTGGDRDIFFG 115
164 ---LKDHNKAA--NYEAMHLKEKENAGGCKTCA-----ILGSTADIGDIVRG 208
116 TDI---SSCRYIKDTSQTKISKLGDQATEKGTHTIDDKKLOEMWTTHGKRIWEGMCA 172
209 LDVWRDINTNKLSEKFOKIFMGGSNRKKO-----NDNNRRKKWKEQRNLITSSNV--260
173 LTNGLSEKKNILODYSYKNLNNAEKDDCCLEKFAFKPOFLRWYVEMSGDFCERKKLE 232
261 -----KHIPKCTCKRHNNFEK-----IPOFLRWLKEWGDCEEGTEV 300
233 DKVEDVICAKDYEGCKNNKSNNSCVKVEYNYITGKKTOYESOGKFNTERKKRPE 292
301 KOLEKIC---ENKNCSEKK---CKNACSSYEKWIKERKNEYLNOSKRFDSDKLTK-K 351
293 YNSISK-KDASEYTLKDK 308
352 NNLYNKPEDSKAVLRSE 368
DB

RESULT 10
US-08-487-826B-6
Sequence 6, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VITAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 749 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-08-487-826B-6

Query Match 15.7%; Score 268.5; DB 2; Length 749;
Best Local Similarity 26.8%; Pred. No. 1.6e-16;
Matches 85; Conservative 45; Mismatches 116; Indels 71; Gaps 15;

6 NEISGCKNPKES--YPMDC-KKNDNSHGACMPPRROKLC-----VRDLTGGELRRP 56
109 SSVFGCKTKISKYKKKKKNCYSNNKYTKPEVCPPRROQLGTYIFLRDNGEG-----163
57 EDILTKFNCACAKETHFAWH-KYKKDNVNAENELKSGKIPGFRKOMYTTGGDRDIFFG 115
164 ---LKDHNKAA--NYEAMHLKEKENAGGCKTCA-----ILGSTADIGDIVRG 208
116 TDI---SSCRYIKDTSQTKISKLGDQATEKGTHTIDDKKLOEMWTTHGKRIWEGMCA 172
209 LDVWRDINTNKLSEKFOKIFMGGSNRKKO-----NDNNRRKKWKEQRNLITSSNV--260
173 LTNGLSEKKNILODYSYKNLNNAEKDDCCLEKFAFKPOFLRWYVEMSGDFCERKKLE 232
261 -----KHIPKCTCKRHNNFEK-----IPOFLRWLKEWGDCEEGTEV 300
233 DKVEDVICAKDYEGCKNNKSNNSCVKVEYNYITGKKTOYESOGKFNTERKKRPE 292
301 KOLEKIC---ENKNCSEKK---CKNACSSYEKWIKERKNEYLNOSKRFDSDKLTK-K 351
DB

QY 293 YNSYSK-KDASEYLKDK 308
Db 352 NNLNFKFEDSKAYLRSE 368

RESULT 11

US-09-210-288-6
; Sequence 6, Application US/09210288
; Patent No. 6392026
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,288
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 749 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
US-09-210-288-6

Query Match 15.7%; Score 268.5; DB 4; Length 749;
Best Local Similarity 26.8%; Pred. NO. 1.6e-16;
Matches 85; Conservative 45; Mismatches 116; Indels 71; Gaps 15;

QY 6 NEISGCPKES-TPDMDC-KKIDNSHSGACMPRRQKLC-----VRDLTGGELRKP 56
Db 109 SSVGCGTKIKISKVKKKMYSNNKVTKEGVCGRPOOLGXYIFILRDNEFG-----163
QY 57 EDILTKFNCIAKETHFAWH-KYKKDYNVAENELKSGKIPGFRKMYTTEGDFRDFPG 115
Db 164 ---LKDHIINKAA--NYEAMHLAKETVENAGDKICNA-----ILGSTADIGDIYRG 208
QY 116 TDI--SSCRVYIKDTSOTIKSLGDQATTEKGDTHIIDNKKLOEWMTIHGPKIWEGMCA 172
Db 209 LDVMDINTNKLSEKFKIFWGGGNSRKKO-----NDNNERNKMWKEQRULIMSSNV-- 260
QY 173 LTNGISESEKKNIILODYSYNKLNAEKDCCLEKFAKSPQFLRWTVENSDFECRRKLE 232
Db 261 -----KHIPKGTCKRRHNEK-----IPQFLRWLKEWGEDEFCEMGTEV 300

QY 293 YNSYSK-KDASEYLKDK 308
Db 301 KOLEKIC-----ENKNCSEKK-----CKNACSEYKEMIKERNKNENYLOSKKIDSDKKLKK-K 351

RESULT 12

US-08-568-459A-21
; Sequence 21, Application US/08568459A
; Patent No. 5849306
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelien, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-568-459A-21

Query Match 15.4%; Score 263.5; DB 2; Length 311;
Best Local Similarity 26.3%; Pred. NO. 1.4e-16;
Matches 67; Conservative 19; Mismatches 158; Indels 11; Gaps 4;

QY 22 CKKIDNSHSGACMPRRQKLC-----VRDLTGGELRKEEDLITKFINCAAEKTHFAWHKY 78
Db 12 CXXXXXXXAXXACMPRRQKLCLYYIXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 71
QY 79 KKDYNVAENELKSGKIPGFRKMYTTEGDFRDIFFGDISCRVYIKDTSOTIKSLGDO 138
Db 72 XXXXXXXXXXXXXXXXXXXXQFLRSMYTTGDIYRDLCLNTDIS-----KKQNVYXXXXXXX 126
QY 139 ATTEKGDTHIIDNKKLOEWMTIHGPKIWEGMCAI--TNGISESEKKNIILODYSYNKLNA 197

Db 127 XXXXXXXXXXXXSPGSLRQEWMTNGPEIWKGMICALXXXXXXXXXXXXXXXXXXXXX 186
QY 198 EKDDCCLEFASKPOFLRWYVWSDFECREKKLEDKVEDVCIRAKDEGCKNNKNSNC 257
Db 187 XXXXXXXXXXXXSKPSGSLRQEWMTNGPEIWKGMICALXXXXXXXXXXXXXXXXXXXXX 244
QY 258 VKCKEYENYITGKK 272
Db 245 NOACRAYOEYVENKK 259

RESULT 13
US-08-487-826B-33
; Sequence 33, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbie Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelisen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-487-826B-33

Query Match 15.4%; Score 263.5; DB 2; Length 311;
Best Local Similarity 26.3%; Pred. No. 1.4e-16;
Matches 67; Conservative 19; Mismatches 158; Indels 11; Gaps 4;

QY 22 CKKNIDNSHSGACMPPRQKLC--VRDLTGGEIRKPEDILTFKFINCAKETHPAMHKY 78
Db 12 CXXXXXXXAXACMPPRQKLCIYIXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 71
QY 79 KKDNNVNAENELSKRIPEGFRKQMYTTFGDFRDIFFGDIDISSCRITKDTISQTIKSKLGDQ 138
Db 72 XXXXXXXXXXXXXXXXXXXXQFLRSMMYTFGDFRDICLMTDIDIS-----KKQNDVXXXXXXXXX 126

QY 139 ATTEGDPHIDNKKLOEWMWTHIGKTIWEGMICAL-TGSLSESEKKNILQDYSYNKLNNA 197
Db 127 XXXXXXXXXXXXSKPSGSLRQEWMTNGPEIWKGMICALXXXXXXXXXXXXXXXXXXXXX 186
QY 198 EKDDCCLEFASKPOFLRWYVWSDFECREKKLEDKVEDVCIRAKDEGCKNNKNSNC 257
Db 187 XXXXXXXXXXXXSKPSGSLRQEWMTNGPEIWKGMICALXXXXXXXXXXXXXXXXXXXXX 244
QY 258 VKCKEYENYITGKK 272
Db 245 NOACRAYOEYVENKK 259

RESULT 14
US-09-210-288-21
; Sequence 21, Application US/09210288
; Patent No. 6392026
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbie Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,288
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-09-210-288-21

Query Match 15.4%; Score 263.5; DB 4; Length 311;
Best Local Similarity 26.3%; Pred. No. 1.4e-16;
Matches 67; Conservative 19; Mismatches 158; Indels 11; Gaps 4;

QY 22 CKKNIDNSHSGACMPPRQKLC--VRDLTGGEIRKPEDILTFKFINCAKETHPAMHKY 78
Db 12 CXXXXXXXAXACMPPRQKLCIYIXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 71
QY 79 KKDNNVNAENELSKRIPEGFRKQMYTTFGDFRDIFFGDIDISSCRITKDTISQTIKSKLGDQ 138

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 15:05:30 ; Search time 32.2489 Seconds
(Without alignments)
1033.453 Million cell updates/sec

Title: US-10-087-013-11

Perfect score: 1706
Sequence: 1 GNGSNDISGCPKESYPDW.....KKPEYNSYSKDASEYLKDK 308

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

1 number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PC7_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PC7US_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|-------------------|
| 1 | 669.5 | 39.2 | 2710 | 9 US-10-153-273-12 | Sequence 12, Appl |
| 2 | 521 | 30.5 | 921 | 9 US-10-153-273-8 | Sequence 8, Appl |
| 3 | 327.5 | 19.2 | 1143 | 10 US-09-924-154-14 | Sequence 14, Appl |
| 4 | 268.5 | 15.7 | 749 | 9 US-10-153-273-6 | Sequence 6, Appl |
| 5 | 268.5 | 15.7 | 1086 | 10 US-09-924-154-15 | Sequence 15, Appl |
| 6 | 263.5 | 15.4 | 311 | 9 US-10-153-273-21 | Sequence 21, Appl |
| 7 | 262.5 | 15.4 | 972 | 10 US-09-924-154-16 | Sequence 16, Appl |
| 8 | 261 | 15.3 | 1421 | 10 US-09-924-154-13 | Sequence 13, Appl |
| 9 | 259 | 15.2 | 1115 | 9 US-10-153-273-2 | Sequence 2, Appl |
| 10 | 258.5 | 15.2 | 1501 | 10 US-09-924-154-17 | Sequence 17, Appl |
| 11 | 251 | 14.7 | 1435 | 9 US-10-153-273-4 | Sequence 4, Appl |
| 12 | 237.5 | 13.9 | 700 | 9 US-10-153-273-10 | Sequence 10, Appl |
| 13 | 221 | 13.0 | 324 | 9 US-10-153-273-17 | Sequence 17, Appl |
| 14 | 156 | 9.1 | 277 | 9 US-10-153-273-15 | Sequence 15, Appl |
| 15 | 135 | 7.9 | 291 | 9 US-10-153-273-13 | Sequence 13, Appl |
| 16 | 127.5 | 7.5 | 282 | 9 US-10-153-273-16 | Sequence 16, Appl |
| 17 | 119.5 | 7.0 | 362 | 9 US-10-153-273-18 | Sequence 18, Appl |
| 18 | 112 | 6.6 | 411 | 9 US-10-153-273-19 | Sequence 19, Appl |
| 19 | 109 | 6.4 | 441 | 9 US-10-153-668-102 | Sequence 102, App |

| | | | | | |
|----|-------|-----|------|-----------------------|-------------------|
| 20 | 109 | 6.4 | 634 | 9 US-10-153-668-98 | Sequence 98, Appl |
| 21 | 109 | 6.4 | 634 | 9 US-10-153-668-100 | Sequence 100, App |
| 22 | 109 | 6.4 | 1118 | 9 US-10-153-668-104 | Sequence 104, App |
| 23 | 102 | 6.0 | 1338 | 10 US-09-402-100-4 | Sequence 4, Appl |
| 24 | 101.5 | 5.9 | 271 | 9 US-10-153-273-14 | Sequence 14, Appl |
| 25 | 101.5 | 5.9 | 497 | 9 US-09-820-843A-32 | Sequence 32, Appl |
| 26 | 101.5 | 5.9 | 665 | 9 US-09-820-843A-107 | Sequence 107, App |
| 27 | 99.5 | 5.8 | 411 | 9 US-10-153-273-20 | Sequence 20, Appl |
| 28 | 97 | 5.7 | 380 | 10 US-09-134-333-13 | Sequence 13, Appl |
| 29 | 94.5 | 5.5 | 526 | 9 US-09-372-348-16 | Sequence 16, Appl |
| 30 | 93.5 | 5.5 | 368 | 10 US-09-925-300-1356 | Sequence 1356, Ap |
| 31 | 93 | 5.5 | 743 | 9 US-10-087-464-53 | Sequence 53, Appl |
| 32 | 93 | 5.5 | 2789 | 10 US-09-801-574-57 | Sequence 57, Appl |
| 33 | 92.5 | 5.4 | 624 | 10 US-09-815-242-5391 | Sequence 5391, Ap |
| 34 | 92.5 | 5.4 | 627 | 10 US-09-815-242-5391 | Sequence 12305, A |
| 35 | 92.5 | 5.4 | 698 | 9 US-10-281-478-2 | Sequence 2, Appl |
| 36 | 92.5 | 5.4 | 698 | 10 US-09-726-968-4 | Sequence 4, Appl |
| 37 | 92.5 | 5.4 | 824 | 9 US-10-281-478-1 | Sequence 1, Appl |
| 38 | 92.5 | 5.4 | 824 | 10 US-09-726-968-2 | Sequence 2, Appl |
| 39 | 92 | 5.4 | 428 | 9 US-09-742-096-3 | Sequence 3, Appl |
| 40 | 92 | 5.4 | 1786 | 9 US-09-824-574-7 | Sequence 7, Appl |
| 41 | 92 | 5.4 | 2476 | 9 US-10-223-085-336 | Sequence 336, App |
| 42 | 91.5 | 5.4 | 776 | 9 US-10-223-084-336 | Sequence 336, App |
| 43 | 91.5 | 5.4 | 776 | 9 US-10-223-088-336 | Sequence 336, App |
| 44 | 91.5 | 5.4 | 776 | 9 US-10-223-090-336 | Sequence 336, App |
| 45 | 91.5 | 5.4 | 776 | 9 US-10-223-090-336 | Sequence 336, App |

ALIGNMENTS

RESULT 1
US-10-153-273-12
Sequence 12, Application US/10153273
Patent No. US20020169305A1
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
Chilnis, Chetan
Miller, Louis H.
Peterson, David S.
Su, Xin-zhaun
Welliams, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/153,273
FILING DATE: 21-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION NUMBER: <Unknown>
APPLICATION NUMBER: US/09/210,288
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121, JFMDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:

LENGTH: 2710 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Plasmodium falciparum
 SEQUENCE DESCRIPTION: SEQ ID NO: 12:
 US-10-153-273-12

Query Match 39.2%; Score 669.5; DB 9; Length 2710;
 Best Local Similarity 40.9%; Pred. No. 8.1e-48;
 Matches 133; Conservative 54; Mismatches 109; Indels 29; Gaps 9;

QY 1 GNDGSEISGNCNPKESYPDMDC-KKNIDNSHGACMPRRKOKL---VRDLTGGEIRKPE 57
 Db 1358 GKNGRITVEGCNPKESYPDMDC-KNNIDISHDGACMPRRKOKLCLYIAHESQENIKITD 1417
 58 DILTKINCAKETHFAHMKYKKNVNAENELKSGKIPGFRKOMYTFGDFRDIFFGTD 117
 1418 NKDAEIKTAAAEPLSMOYKSKNDSEAKILDRGLIPSOFLSMYTFGDYRDICLND 1477
 QY 118 ISSCRYIKDTSOTIKSLG----DQATEKGDTHTDDKKLOEMWTIGPKTWEGMLCA 172
 Db 1478 ISKQ--NDYAKA-KKIGKFSKDSKSGSLR-----QEWKTPGPELWKMGLCA 1527
 QY 173 LINGLSESE-KNNIIDODYSYNKLNNAEKDDCCLEKFPASRPOFLRWYVENSDFECPRRKT 231
 Db 1528 LKRYVDTDNKRKIKNDYSYKVNQNGNPISLEFAAPQFLRWIEMGEFCAROK 1587
 QY 232 EKVVEYCIKADYDECKNNKSNNSCVKCKEYENITGKTKTOYESQESKFN--NTEKKRK 289
 Db 1588 ENIADACNEINSTOCC--NDAKHRCNQACRAOEVYENKKKEFSQTNINFLKANVQPO 1645
 QY 290 KPEYNSYSRKD-----ASEYLKDK 308
 Db 1646 DEYKGYEYKDGVOPIQGENEYLLOK 1670

RESULT 2

US-10-153-273-8
 Sequence 8, Application US/10153273
 Patent No. US20020169305A1
 GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
 Chitnis, Chetan
 Miller, Louis H.
 Peterson, David S.
 Su, Xin-zhaun
 Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/153.273
 FILING DATE: 21-May-2002
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/210.288
 FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
 NAME: Fuller, Michael
 REGISTRATION NUMBER: 36,516
 REFERENCE/DOCKET NUMBER: NIH21.1FMDV1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 921 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Plasmodium falciparum
 SEQUENCE DESCRIPTION: SEQ ID NO: 8:
 US-10-153-273-8

Query Match 30.5%; Score 521; DB 9; Length 921;
 Best Local Similarity 37.9%; Pred. No. 8.3e-36;
 Matches 127; Conservative 46; Mismatches 106; Indels 56; Gaps 13;

QY 3 DGSNEISGNCNPKESYPDMDC-KKNIDNSHGACMPRRKOKLYRDLTGGEIRKPELIT 61
 Db 407 NGRTIVGECIKREITSEMTCDSEKTKMGCHGACIPRRKOKLHYL-----EKIMT 457
 QY 62 K-----FINCAKETHFAHMKYKRD--NVNAEN--ELKSGKIPGFRKOMYTFGDFR 110
 Db 458 NTELKVAEICAIAEFTELWQNYKRDKNAGNAEDLDEKLKGIIPEDKRQMFETPADYR 517
 QY 111 DIFPGTSSCRYIKDTSOTIKSLGQDATERKGTHTD-----NKKLOEMWTIG 162
 Db 518 DICLGTDIS--KKOTISKV-----GKVKCNIDVFIKINSIRYKSWMETNG 564
 QY 163 PRWEGMLCALTLNGLS-----ESEKNNIIDODYSYNKLNNAEKDDCCLEKFPASRPOFLRW 216
 Db 565 PVWEGMLCALSYDPLSNVNPETHKKLTEBNNNFYKIPESDSTLSKSEKPPOLRW 624
 QY 217 YVENSDFECPRRKLEKVEYDVCIKADYDECKNNKSNNSCVK--CKEYENITGKKT 273
 Db 625 LTEWGENECKBOKK-EYKV--LLAKCKDQVDGDKGCKNGKCVACKDCKYHSMIGIMID 681
 QY 274 QYESQEGKFNTERK--QKKPEYNSYSRKDDASEYLK 306
 Db 682 NYKKQKGRYEVKKIPLKEDKDVKNSDDARDYLK 716

RESULT 3

US-09-924-154-14
 Sequence 14, Application US/09924154
 Patent No. US20020127241A1
 GENERAL INFORMATION:

APPLICANT: Narum, David L.
 TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use

FILE REFERENCE: 05213-0465 43170-262105
 CURRENT APPLICATION NUMBER: US/09/924.154

PRIOR FILING DATE: 2001-08-07
 PRIOR APPLICATION NUMBER: US 60/223,525

PRIOR FILING DATE: 2000-08-07
 NUMBER OF SEQ ID NOS: 17

SOFTWARE: Patentin version 3.1
 SEQ ID NO 14

LENGTH: 1143
 TYPE: PRT

ORGANISM: Mammalian
 US-09-924-154-14

Query Match 19.2%; Score 327.5; DB 10; Length 1143;
 Best Local Similarity 27.8%; Pred. No. 2.8e-19;
 Matches 85; Conservative 46; Mismatches 118; Indels 57; Gaps 10;

```
QY 10 GCNPK-ESYDPMDCCKNIDNSHSGACMPPRROKLCVDR---LNGGGEIRKPEDILTFIN 65
D 461 GCNKSNNKSNKSWNCTGFTTKPFCPTCEPPRROTLCGRITLLHGHNE---EDKKEHLG 516
QY 66 CAKETHFAHMKYKKKDVNAENELKSGKIEGFRKQMYTFGDFRDIFFGDTLSSCRVYK 125
D 517 ASIYEAQLKYYKKEKDNALCSI-----IQNSYADLADILKSDIIDLKDYK 564
QY 126 DTSGTTSKLGDAQATTEKGDTHIDNNKL--QEWTHIGPKIWEGLCALTNLSESEK 183
D 565 KMEENLKNVKKDKRNE-----ESLKIFREKWMDEKNKENVKMSAVLKN-----K 610
QY 184 NIIDYSYNKLNNAEKDDCCLEKFAKPOFLRWYVEMSDPCREBRKLEDEVCITAK 243
D 611 ETCDY-----DKFOKITPFLRMFKKNGDPCERKKKITYSFESFVECK 655
QY 244 DYECCKNNKSNKSVCKVEYENYITGKTQYESQEGKFNTERKQK-PEYNSYSKKDAS 302
D 656 ----KDCDENTCKNCKSEYKKWIDLKSEYKQVDRKTKDKNKNYDNIDEVKNKEAN 710
303 EYLKDK 308.
1111
D 711 VYLK 716

RESULT 4
US-10-153-273-6
: Sequence 6, Application US/10153273
: Patent No. US20020169305A1
: GENERAL INFORMATION:
: APPLICANT: Sim, Kim L.
: Chitnis, Chetan
: Miller, Louis H.
: Peterson, David S.
: Su, Xin-zhaun
: Wellem, Thomas E.
: TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Knobbe Martens Olson & Bear
: STREET: 620 Newport Center Drive 16th floor
: CITY: Newport Beach
: STATE: California
: COUNTRY: US
: ZIP: 92660
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/153,273
: FILING DATE: 21-May-2002
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/09/210,288
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Fuller, Michael
: REGISTRATION NUMBER: 36,516
: REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 235-8550
: TELEFAX: (619) 235-0176
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 749 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
```

```
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-153-273-6
Query Match 15.7%; Score 268.5; DB 10; Length 749;
Best Local Similarity 26.8%; Pred. No. 1.7e-14;
Matches 85; Conservative 45; Mismatches 116; Indels 71; Gaps 15;

QY 6 NEISGCNPKES--YPMDC-CKNIDNSHSGACMPPRROKLC-----VRDLTGGELTRKP 56
D 109 SSVFGCKTKISKYKKKNKCYNNKVTPEGYCGPPRQQLGLGIFILRGNNEG----- 163
QY 57 EDILTFINCAAKETHFAH- KYKKNVNAENELKSGKIEGFRKQMYTFGDFRDIFFG 115
D 164 ---LKHINKAA--NYEAMHLKEKYENAGDKICNA-----ILGSYADIGDIVRG 208
QY 116 TDI---SSCRVIRKDSQTSKLGDAQATTEKGDTHIDNNKLQEWTHIGPKIWEGLCA 172
D 209 LDVWRDINTNKLSEKPKITFMGGNSRRKQ-----NDNNEKKNWKEKORNLWSSNV-- 260
QY 173 LFNGLSESEKKNIIIDYSYNKLNNAEKDDCCLEKFAKPOFLRWYVEMSDPCREKRL 232
D 261 -----KHIPKGTCKRHHNFEK-----IPQFLRWLKEGDEFCCEMCTEV 300
QY 233 DKVEDVCITAKDYGCKNNKSNKSVCKVEYENYITGKTQYESQEGKFNTERKQKPE 292
D 301 KLEKIC---ENKNCSEK---CKNACSSYEYKIKERKNENYLSQKFDSDKLNK-K 351
QY 293 YNSYSK-KDASEYLLDK 308
D 352 NNLNKEFEDSKAVYLRSE 368

RESULT 5
US-09-924-154-15
: Sequence 15, Application US/09924154
: Patent No. US20020127241A1
: GENERAL INFORMATION:
: APPLICANT: Narum, David L.
: APPLICANT: Sim, Kim L.
: TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
: FILE REFERENCE: 05213-0465 43170-262105
: CURRENT APPLICATION NUMBER: US/09/924,154
: CURRENT FILING DATE: 2001-08-07
: PRIOR APPLICATION NUMBER: US 60/223,525
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: Patentln version 3.1
: SEQ ID NO 15
: LENGTH: 1086
: TYPE: PRT
: ORGANISM: Mammalian
US-09-924-154-15
Query Match 15.7%; Score 268.5; DB 10; Length 1086;
Best Local Similarity 26.8%; Pred. No. 2.7e-14;
Matches 85; Conservative 45; Mismatches 116; Indels 71; Gaps 15;

QY 6 NEISGCNPKES--YPMDC-CKNIDNSHSGACMPPRROKLC-----VRDLTGGELTRKP 56
D 351 SSVFGCKTKISKYKKKNKCYNNKVTPEGYCGPPRQQLGLGIFILRGNNEG----- 405
QY 57 EDILTFINCAAKETHFAH- KYKKNVNAENELKSGKIEGFRKQMYTFGDFRDIFFG 115
D 406 ---LKHINKAA--NYEAMHLKEKYENAGDKICNA-----ILGSYADIGDIVRG 450
QY 116 TDI---SSCRVIRKDSQTSKLGDAQATTEKGDTHIDNNKLQEWTHIGPKIWEGLCA 172
D 451 LDVWRDINTNKLSEKPKITFMGGNSRRKQ-----NDNNEKKNWKEKORNLWSSNV-- 502
QY 173 LFNGLSESEKKNIIIDYSYNKLNNAEKDDCCLEKFAKPOFLRWYVEMSDPCREKRL 232
```

Db 503 -----KHPKGTCKTCKRHNNFEK-----IPQFLRWLKMEDDFCEENGTEV 542
QY 233 DKVEDVICAKDGECKNNKSNNSCYKVCKEYENITYTGGKTOYESOEGFNTKROKPE 292
Db 543 KQLEKIC---EKKNKSEKK---CKNACSYEKWKERKNEYNLOSKKFDSDKLNK-K 593
QY 293 YNSYSK-KDASEYLKDK 308
Db 594 NNLNKFEDSKAYLRSE 610

RESULT 6
US-10-153-273-21
; Sequence 21, Application US/10153273
; Patent No. US20020169305A1
; GENERAL INFORMATION:
; APPLICANT: Slim, Kim L.
; Chitnis, Chetan
; Miller, Louis H.
; Peterson, David S.
; Su, Xin-zhaun
; Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/153,273
; FILING DATE: 21-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,288
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION: }
; NAME: Fullier, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-153-273-21

Query Match 15.4%; Score 263.5; DB 9; Length 311;
Best Local Similarity 26.3%; Pred. No. 1.5e-14;
Matches 67; Conservative 19; Mismatches 158; Indels 11; Gaps 4;

QY 22 CKKNIDNSHGACMPRRKLC---VRDLTQGEIRKPEDILTRKFINCAKETHFAMHKY 78
Db 12 CXXXXXXXXXXACMPRRKLCIYIIXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 71

QY 79 KQDVNAENELKSGIPEGFRKQMYTFEGDFRDIFFGTDISCRYIKDTQSOTIKSLDQ 138
Db 72 XXXXXXXXXXXXXXXXXXXXQFLRSMYTFGDRDILCLNDIS-----KQNDVYXXXXXXX 126
QY 139 ATTERGDIHIDNKKLOEWMTIHGKTIWEGMLCAL-TNGLSESEKKNILQDYTNKLNNA 197
Db 127 XXXXXXXXXXXXSPSGLSRQEMWMTNGPELWGMICALXXXXXXXXXXXXXXXXXXXXXXX 186
QY 198 EKDDCLEKFAKRPQFLRWYVWSDPFCRERKRLKEDKVEDVICAKDVEGCKNNKSNNSC 257
Db 187 XXXXXXXXXXXXXKQFLRWMTLMEWGEBCAFERKKNENIKDQXXXXXXXXXXCXKXK--HRC 244
QY 258 KVCKEYENITYGKK 272
Db 245 NQACRAVOEYENRK 259

RESULT 7
US-09-924-154-16
; Sequence 16, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:
; APPLICANT: Narum, David L.
; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
; FILE REFERENCE: 05213-0465 43170-262105
; CURRENT FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US/09/924,154
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Version 3.1
; SEQ ID NO 16
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Mammalian
US-09-924-154-16

Query Match 15.4%; Score 262.5; DB 10; Length 972;
Best Local Similarity 24.7%; Pred. No. 7.5e-14;
Matches 78; Conservative 51; Mismatches 130; Indels 57; Gaps 11;

QY 11 CNPESYPMDDCKKNIDNSHGACMPRRKLCVRLDTGGELIRKEDILTRKIN-----C 66
Db 138 CKKGNKTIWQCTNEHKKDPDYCGPFRQDCLGNDK-DEKKNVND-LKKFLNETITIG 195
QY 67 AAKETHFAMHKYKQDVNAENELKSGIPEGFRKQMYTFEGDFRDIFFGT-----IS 119
Db 196 IRDEGKFLIEKRYK-----NMHENWYLDERACKYLYNSPDYKKNILIGKDMWRDPSNIR 249
QY 120 SCRYIKDTQSOTIKSLDQATTEKGTTHIDNKKLOEWMTIHGPKIWEGLCALINGLSE 179
Db 250 TENILKGNFEGILANIVSMYPS-YADLSIDEPFR--HMDQNNKQMLEAISEFYGK-- 303
QY 180 SEKKNILQDYSYNKLNNAEKDCLKFAKRPQFLRWYVWSDPFCRERKRLKEDKVEDVC 239
Db 304 -----NHTGVLMBDDNDNGLMFEMKKNDFCIDLKNNDYKKEC 345
QY 240 IAKKDYEGCKNNKSNNS-----CVKCKEYENITYGKKTQYESOEGFNTKROKPE 292
Db 346 IDRK---VKSPPSPENPSDVATVCKSCTDYDKMIINRKREYKMQSSKY---KRDRSLF 398
QY 293 YNSYSKKDASEYLKDK 308
Db 399 NNVIONLKPWEYLSMK 414

RESULT 8
US-09-924-154-13
; Sequence 13, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:


```

: APPLICANT: Narum, David L.
: APPLICANT: Sim, Kim L.
: TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
: FILE REFERENCE: 05213-0465 43170-262105
: CURRENT APPLICATION NUMBER: US/09/924,154
: CURRENT FILING DATE: 2001-08-07
: PRIORITY APPLICATION NUMBER: US 60/223,525
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 13
: LENGTH: 1421
: TYPE: PRT
: ORGANISM: Mammalian
: US-09-924-154-13

```

Query Match 15.38; Score 261; DB 10; Length 1421;

Best Local Similarity 27.28; Pred. No. 1.6e-13;

Matches 81; Conservative 46; Mismatches 101; Indels 70; Gaps 14;

```

20 WDCKKNIDNSHSGACMPRRKOKLCVRDLTGGGIRKPED-----ILKRFICAKETHFA 74
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 474 WECKKPKLSTKDYCVPRROELCL-----GNIDRIYDKNLMITKEHILAIAIYESRIL 527
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 75 WHKXKDNVNAENELKSGKIPGFRKQMYTFGDFRDIFFGTDISCRITKDTISQITISK 134
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 528 KRXKX-----NKDKEVCKITNKTFADRIDIGTD-----YWNDSL---NRK 567
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 135 LGDQATTEKGDTHID-DNKKL--OEWMTIHGPKIWEGMLCALITNGISESEKKNILQDYSY 191
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 568 LVGCTININSYVHNKQNDKLFDEMWKVIKKDW-----NVISWVEK----- 610
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 192 NKLNAKDDCCLEFASKPOFLRWYVWMSDFECRERKKLEKVEDYCIKAKDEGGKNN 251
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 611 -----DKYCKEDIEDINIQFFRFESEMGDDYQDXTKM---IETLKECKE-KPEE-- 658
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 252 KSNNSCVKCEYENYITGKTQYESQGFNTEKROKKPEYNSYKRDASE---YTK 306
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 659 --DNCCKRKCNSYKEMISKKEEYKQAKQY--OEOYKGNMYKMSSEKSIKPEVYTK 712
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

RESULT 9

US-10-153-273-2

Sequence 2, Application US/10153273

Patent No. US20020169305A1

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

Chitnis, Chetan

Miller, Louis H.

Peterson, David S.

Su, Xin-zhaun

Wellems, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knodbe Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/153,273

FILING DATE: 21-MAY-2002

CLASSIFICATION: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Fuller, Michael

REGISTRATION NUMBER: 36,516

REFERENCE/DOCKET NUMBER: NIH121.1PMDV1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1115 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORGANISM: Plasmodium vivax

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-153-273-2

Query Match 15.28; Score 259; DB 9; Length 1115;

Best Local Similarity 24.98; Pred. No. 1.8e-13;

Matches 89; Conservative 38; Mismatches 127; Indels 104; Gaps 14;

```

QY 1 GNDGSNETSGCNP-----KESYPMDCCKNIDNS 29
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 223 GNSRKNSSNGSNPFDIDHKTITSSAIINHAFLQTVAKNCYKRRERDMDC-----NT 277
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 30 HSGACMPRRKOKLCVRDLT--OGGIRKPEDIL-----TKFICAKETHFAW---- 75
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 278 KDVCIIDPRKQYOLCMKELTNLVNNTDINFRHDIKRLKLYDAVAEGDILLKNN 337
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 76 HKYKKNVNAENELKSGKIPGFRKQMYTFGDFRDIFFGTDISCRITKDTISQITISK 135
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 338 YRYNKD-----FCKDIRWSLGDGDIIMGDMGIGSKYVENNLSIF 381
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 136 GDQATTEKGDTHIDNKKLOEWMTHGPKIWEGMLCALITNGISESEKKNILQDYSYKLN 195
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 382 G---TDEKA-----QGRKQWMSKAKQITAMYSYKRRL---KGNFIWICKLNAV 428
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 196 NAEKDDCCLEFASKPOFLRWYVWMSDFECRERKKLEKVEDYV---IKADYEGCKNNK 252
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 429 NIE-----POLYRWIRMGGRDYSELPTEYOKLKEKDGKINTYDKRVCK--- 473
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 253 SNNSCVKCEYENYITGKTQYESQGEF-----NTEKROKKPEYNSY--KRDASEY 304
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 474 -VPCQNAKSYDQWITTKKNQMDVLSNKFISVNAEKVQTAGIVTPYDILKQELDEF 530
   |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

RESULT 10

US-09-924-154-17

Sequence 17, Application US/09924154

Patent No. US20020127241A1

GENERAL INFORMATION:

APPLICANT: Narum, David L.

TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use

FILE REFERENCE: 05213-0465 43170-262105

CURRENT APPLICATION NUMBER: US/09/924,154

PRIOR FILING DATE: 2001-08-07

PRIOR APPLICATION NUMBER: US 60/223,525

NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn version 3.1

SEQ ID NO 17

LENGTH: 1501

TYPE: PRT

ORGANISM: Mammalian

US-09-924-154-17

Query Match 15.28; Score 258.5; DB 10; Length 1501;

Best Local Similarity 23.4%; Pred. No. 2.8e-13;

Matches 73; Conservative 61; Mismatches 119; Indels 59; Gaps 13;

QY 11 CNPKRSYPMDOCKRIINDS--HSGACMPPRROKLCVRLDLO--GGEIRKPED--ILTFE 64
 Db 435 CLEKEFPKCDKNKSFETVHHKGVCSPPRQFGGLNLVNDLNDIIVNNSQLLEII 494
 QY 65 NCAKETPHFAHKKYK--DNVNAENELKSGKIPGFRKQMYTGTGDFRIDFGDI---- 118
 Db 495 MASQEGKLLMKKHTIILDNONA-----CKYINDSYVDYKDIYIGNDLWMDN 541
 QY 119 SSCRIKDTQSITKSLGDQATTEKGDTHIDNKKLOEMWTIHGPKIMEGMLCALNGLS 178
 Db 542 NSIVQNNLNLIFERNGVKVRNKLRTIKELKNV--WILNKKYVESNRC---GID 595
 QY 179 ESEKKNLQDYSYKLNNAEKDDCCLEKFAKSPQFLRWYVWSDFCRERKKLEDEYDV 238
 Db 596 EVDPR-----KTCER---IDELNMPQFFWFSQMAHFCEKEKEYELKNDK 641
 QY 239 CIRAKDYEGCKNNKS--NNSCVKCKEYENYITGKKTQYESQEGKNTTERKQKPEYNS 295
 Db 642 CTG-----NNGSKLQDQKTCQNVCTNNMWTYTRKRLAYELQSVKDKDKRL---FSL 690
 QY 296 YSKKDASEYIKD 307
 Db 691 AKDKNVTFLKE 702

RESULT 11

US-10-153-273-4
 ; Sequence 4, Application US/10153273
 ; Patent No. US20020169305A1
 ; GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

Chilnis, Chetan
 Miller, Louis H.
 Peterson, David S.
 Su, Xin-zhaun
 Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/153,273
 FILING DATE: 21-May-2002
 CLASSIFICATION: <unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/210,288
 FILING DATE: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Fuller, Michael
 REGISTRATION NUMBER: 36,516
 REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1435 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Plasmodium falciparum

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-153-273-4

Query Match 14.7%; Score 251; DB 9; Length 1435;

Best Local Similarity 26.3%; Pred. No. 1,1e-12;

Matches 79; Conservative 47; Mismatches 100; Indels 74; Gaps 14;

QY 20 WDCCKNIDNSHSCACPPRRQKLCVRLDLOGGIRKPED--ILTFKINCAAEETHA 74
 Db 474 WECKNYILSTKVCYCPPRQELC-----GNIDRIYKNLMLIKETILALAIYESRL 527
 QY 75 WHKRYDNVNAENELKSGKIPGFRKQMYTGTGDFRIDFGDISSCRIFKPTQISK 134
 Db 528 KRYK-----NKDKREVCKIINKTFADIRDIIGTD-----YWNLSN--RKL 568
 QY 135 LGDQATTEKGDTHIDNKK-----LOEMWTIHGPKIMEGMLCALNGLSESEKKNILQDY 189
 Db 569 VKKINTNSK--YVHRNKKNDKLFREDWKKVIRKQDV-----NVISWVK----- 610
 QY 190 SYNKLNNAEKDDCCLEKFAKSPQFLRWYVWSDFCRERKKLEDEYDVCIKADYEGCK 249
 Db 611 -----DKTVCKEDDLENIPIQFFRWFSQMGDDYQDQKTM--ITLKVECKE-KPCE 658
 QY 250 NKSNNSCYKCKEYENYITGKKTQYESQEGKNTTERKQKPEYNSYKSKDASE--YLK 306
 Db 659 ----DNDCKSKNSYKEMISKKEEYKQAKQY--QFYQKGNVYKMYSEFSINREYVLK 712

RESULT 12

US-10-153-273-10
 ; Sequence 10, Application US/10153273
 ; Patent No. US20020169305A1
 ; GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

Chilnis, Chetan
 Miller, Louis H.
 Peterson, David S.
 Su, Xin-zhaun
 Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/153,273
 FILING DATE: 21-May-2002
 CLASSIFICATION: <unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/210,288
 FILING DATE: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Fuller, Michael
 REGISTRATION NUMBER: 36,516
 REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1435 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

```

      FILING DATE: <Unknown>
      ATTORNEY/AGENT INFORMATION:
        NAME: Fuller, Michael
        REGISTRATION NUMBER: 36,516
        REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
        TELECOMMUNICATION INFORMATION:
          TELEPHONE: (619) 235-8550
          TELEFAX: (619) 235-0176
      INFORMATION FOR SEQ ID NO: 17:
        SEQUENCE CHARACTERISTICS:
          LENGTH: 324 amino acids
          TYPE: amino acid
          STRANDEDNESS: single
          TOPOLOGY: linear
          MOLECULE TYPE: peptide
          HYPOTHETICAL: NO
          ANTI-SENSE: NO
          FRAGMENT TYPE: Internal
          ORIGINAL SOURCE:
            SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-153-273-17

Query Match 13.0%; Score 221; DB 9; Length 324;
Best Local Similarity 25.5%; Pred. No. 6,3e-11;
Matches 69; Conservative 17; Mismatches 119; Indels 66; Gaps 8;

QY 33 ACMPRRKQICVNDL--TGGELRKREDILTFINCAKETFPAMHKKKDQVNAENELK 90
Db 24 ACIPRRQKCLHYLAXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 83
QY 91 SGKIPGFKROMYVTFGDFRDIFFGDISSCRYIKDTS-----QTIKSLGDOAT 140
Db 84 XXXXXXDFKRMQYETPADYRDICLGDIS--KKDTSXXXXXXXXXXXXXXXXXISNR 140
QY 141 TEGSDPHIDNKKLQEWMTIHGKRIWEGMLCAL-----TNGLSSEKKNIIDYSYKL 194
Db 141 YRK-----SWETNGFVIEGMLCALXXXXXXXXXXXXXXXXXXXXXXXXXXXX 188
QY 195 NNAEKDDCLEKFAKPOELRMVVEWSDFCERKRKLIEDKV-----ED 237
Db 189 XXXXXXXXXXXXXXXXRPQFLRMLETGENFCKEQKK-EYKVLACXXXXXXXXXXXX 247
QY 238 VCIKADYEGCKNNKSNNSCVAKCYEYNTI 268
Db 248 XCVACKDQ-----CKQYHSMI 263

RESULT 14
US-10-153-273-15
: Sequence 15, Application US/10153273
: Patent No. US20020169305A1
GENERAL INFORMATION:
APPLICANT: Slim, Kim L.
Chilnits, Chetan
Miller, Louis H.
Peterson, David S.
Su, Xia-zhaun
Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knodde Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```

[illegible]

| | | | | | | |
|----|---|--------|-------------------|-------|-------------|-----|
| | Query Match | 7.9%; | Score 135; | DB 9; | Length 291; | |
| | Best Local Similarity | 22.2%; | Pred. No. 0.0011; | | | |
| | Matches | 41; | Conservative | 19; | Mismatches | 79; |
| | | | | | Indels | 46; |
| | | | | | Gaps | 6 |
| QY | 98 FRKQMTYTGGDPDRDIFFGTIDISCCRYIKDTQSOTIKSKIGDAQTEKEGDTHTDDNKKIQDEM | 157 | | | | |
| Db | 83 FCDIDIMWSJGDFGDIDIMGDMGISYRXXXXXXX--XXXTDEKA-----QQRRKW | 133 | | | | |
| QY | 158 WTTHGEKRIWEGMICALTNGLSESEKKNILODYSYNKRLNNNAEKDDCC--LEKFASKPQFLR | 215 | | | | |
| Db | 134 WNSKRQIWTAM-----YSVAXXXXXXXXXXKCXXXXXXXXXEPQIYR | 175 | | | | |
| QY | 216 WYWEWSDEFCECRERKKLEDVEDVCITKADYES-----CKNNKSNNSCVKVCKEYENT | 267 | | | | |
| Db | 176 WIREWGRDYSELPEVOKLKERKCYAXXXXXXXXXXCXVPQN-----ACKSYDW | 226 | | | | |
| QY | 268 ITGKK | 272 | | | | |
| Db | 227 ITRKK | 231 | | | | |

```
Search completed: June 20, 2003, 15:18:23
Job time : 33.2489 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 14:59:06 ; Search time 24.3035 Seconds

(without alignments)
1218.319 Million cell updates/sec

Title: US-10-087-013-11

Perfect score: 1706

Sequence: 1 GNDGSENEISGNCMPKESYDPM.....KKREYNSYSKKDASEYLKDK 308

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Minimum number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|--------------------|
| 1 | 894 | 52.4 | 3006 | 2 | T28625 | variant-specific s |
| 2 | 669.5 | 39.2 | 3078 | 2 | T28432 | variant-specific s |
| 3 | 600 | 35.2 | 3026 | 2 | T28431 | variant surface pr |
| 4 | 526 | 30.8 | 2706 | 2 | T28155 | variant-specific s |
| 5 | 433 | 25.4 | 1711 | 2 | C71625 | variant-specific s |
| 6 | 395 | 23.2 | 2042 | 2 | T18399 | variant-specific s |
| 7 | 385 | 22.6 | 2228 | 2 | T14029 | variant-specific s |
| 8 | 383 | 22.5 | 2212 | 2 | T28157 | erythrocyte membra |
| 9 | 367.5 | 21.5 | 2182 | 2 | T28634 | variant-specific s |
| 10 | 361.5 | 21.2 | 2647 | 2 | T28161 | hypothetical prote |
| 11 | 354 | 20.8 | 2135 | 2 | T14602 | variant-specific s |
| 12 | 352 | 20.6 | 2664 | 2 | T28626 | variant-specific s |
| 13 | 342.5 | 20.1 | 2197 | 2 | B71600 | variant-specific s |
| 14 | 311 | 18.2 | 2924 | 2 | T18378 | variant-specific s |
| 15 | 298.5 | 17.5 | 1729 | 2 | T18396 | erythrocyte membra |
| 16 | 259 | 15.2 | 1070 | 2 | T30848 | Duffy receptor - p |
| 17 | 251 | 14.7 | 1435 | 2 | A37793 | erythrocyte-bindin |
| 18 | 246.5 | 14.4 | 1153 | 2 | T28652 | erythrocyte-bindin |
| 19 | 237.5 | 13.3 | 1045 | 2 | T18373 | erythrocyte-bindin |
| 20 | 178 | 10.4 | 778 | 2 | A35970 | erythrocyte-bindin |
| 21 | 121 | 7.1 | 2523 | 2 | T18477 | hypothetical prote |
| 22 | 117 | 6.9 | 1282 | 2 | JE0120 | glycoprotein A - m |
| 23 | 114.5 | 6.7 | 508 | 2 | E71620 | hypothetical prote |
| 24 | 113.5 | 6.7 | 763 | 2 | G97026 | superfamily I DNA |
| 25 | 111.5 | 6.5 | 696 | 2 | G71620 | hypothetical prote |
| 26 | 110.5 | 6.5 | 1712 | 2 | C71618 | hypothetical prote |
| 27 | 110.5 | 6.5 | 3724 | 2 | T18427 | hypothetical prote |
| 28 | 109 | 6.4 | 1939 | 2 | T18372 | repeat organellar |
| 29 | 109 | 6.4 | 2391 | 2 | T18410 | carbamoyl-phosphat |

| | | | | | | |
|----|-------|-----|------|---|--------|--------------------|
| 30 | 108 | 6.3 | 433 | 2 | T25946 | hypothetical prote |
| 31 | 107.5 | 6.3 | 219 | 2 | T19897 | hypothetical prote |
| 32 | 107 | 6.3 | 1997 | 2 | F71607 | DNA helicase II BR |
| 33 | 106.5 | 6.2 | 463 | 2 | S27757 | embryonic abundant |
| 34 | 106.5 | 6.2 | 3844 | 2 | T18402 | asparagine/asparta |
| 35 | 106 | 6.2 | 1138 | 2 | E71615 | probable amine tra |
| 36 | 105.5 | 6.2 | 446 | 2 | T08187 | hypothetical prote |
| 37 | 105 | 6.2 | 522 | 2 | C96608 | hypothetical prote |
| 38 | 105 | 6.2 | 1302 | 1 | JC6009 | surface-located me |
| 39 | 103 | 6.0 | 342 | 2 | T18474 | hypothetical prote |
| 40 | 103 | 6.0 | 479 | 2 | T47561 | late embryogenesis |
| 41 | 103 | 6.0 | 807 | 2 | T18454 | hypothetical prote |
| 42 | 103 | 6.0 | 930 | 2 | D71617 | SERA antigen/papal |
| 43 | 103 | 6.0 | 1025 | 2 | S69790 | fibronectin-bindin |
| 44 | 103 | 6.0 | 1192 | 2 | A71623 | probable secreted |
| 45 | 103 | 6.0 | 1855 | 2 | S41649 | DNA polymerase - m |

ALIGNMENTS

RESULT 1

T28625

variant-specific surface protein 3 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T28625

R:Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herrfeldt, J.A.; Peterson, Cell 82, 89-100, 1995

A:Title: The large diverse gene family var encodes proteins involved in cytoadherence

A:Reference number: 220487; MUID:95330813; PMID:7606788

A:Accession: T28625

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3006 <SUX>

A:Cross-references: EMBL:LA0609; NID:9886376; PID:9886377; PIDN:AAA75397.1

C:Genetics:

A:Gene: var-3

A:Introns: 2597/3

| | | | | |
|-----------------------|------------------|---|------------|--------------|
| Query Match | 52.4% | Score 894; | DB 2; | Length 3006; |
| Best Local Similarity | 55.0% | Pred. No. 4.3e-56; | | |
| Matches 172; | Conservative 38; | Mismatches 91; | Indels 12; | Gaps 4; |
| QY | 5 | SNEISGNCNP-----KESYPMDCCKKNINDNSHSGCMPPRRQKLCVRLDTGGELRRKEDI | 59 | |
| DB | 1236 | TDDIDCGNOKYKAGKDKYPMDCNSOIHITHNGACMPRRQKLCVSLGFTDRKALEYI | 1295 | |
| QY | 60 | LTKFINCAAKETHPAWHKRYKKNVNAENELKSGKIPGCFRKNMYTFGDFRDIFFGTDIS | 119 | |
| DB | 1296 | RTEFIKSALETHPAWHRYKEDNGEALKNGNIPGCFRKNMYTFGDFRDIFFGTDIS | 1355 | |
| QY | 120 | SCRYIKDTSQ--TIKSLGDQATTEKGDTHIDNNKRLQEWMTIHGPKIWEIGMICALTNG | 176 | |
| DB | 1356 | THAIVISGVSFKVITILEKENDAKYAAKONS---NNEILLDDWMDHGDIDIEGMICALTHK | 1412 | |
| QY | 177 | LSSESE-KKNIIQDYSYKKNLNNAEKDCELEFAKPPQLFWYKMSDFEPRERKKLEDDKY | 235 | |
| DB | 1413 | ISDEKKRKETKKNYSYKKNLSPGNSKRVDFEFAKPPQLFWYKMSDFEPRERKKLEDDKY | 1472 | |
| QY | 236 | EDVCIRAKDEGCKNNNSNCVAVCKEYENYINGKTKQYESDQKFNTEKROKPPYNS | 295 | |
| DB | 1473 | KVSCSDAKDVGDCGNTNMSNACVACVIEDYITKKKVEYTKQKGRDAEKITDKBEYEG | 1532 | |
| QY | 296 | YSKKDASEYLKDK 308 | | |
| DB | 1533 | FSTKDASEYLKDK 1545 | | |
| RESULT 2 | | | | |
| T28432 | | variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum) | | |
| N:Alternate names: | | erythrocyte membrane binding protein 1 (EMP1) | | |

C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: C71625
R:Gardner, M.J.; Tetteh, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
I:Partea, M.; Salzberg, S.; Zhou, L.; Sutton, R.; White, O.; Smith, H.O.
Science 282, 1126-1137, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference numbers: A11600; MUID:99021743; PMID:9804551
A:Accession: C71625
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1711 <GAR>
A:Cross-references: GB:A0001366; GB:AE001362; NID:g3845070; PIDN:AAC71792.1; PID:g3845070
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0010W

| | | | | |
|-----------------------|------------------|--------------------|------------|--------------|
| Query Match | 25.4% | Score 433; | DB 2; | Length 1711; |
| Best Local Similarity | 33.4% | Pred. No. 3.8e-23; | | |
| Matches 104; | Conservative 49; | Mismatches 110; | Indels 48; | Gaps 15 |

```

20  WDC - KKNIDN$HSGA - CMPRPRRLK - - - - - VRDTCGCEIRKPEDITTKFKCAKE 70
    | - - - - - | | | | | | | | | | | | | | | | | | | | | | | | | | | |
856 WRCIAPSGTSGKQCAICVPRPTQELCYLLKELSDTTQKG - - - - - LREAFITTAQE 908

```

```
QY      71THEFAHMYKRDVN-----AENELKSQKPIEGGRKOMYTFEGRDIFCTDIS 119
        :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :
Db      909 TYLLMOKVYKKEDKONETASTELDIDDPOTQLNGEIPEDFKROMFYTGDRDLFLG --- 964
```

```

120 SCRINIVISDIINSKAGDATTESGTRIDUNKAL--QEWMTIHGPKIWEIGMICALIN I/5
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
965 --RTIGNDELDKVNNNI--TAVFONG-HIPNGOKTRDRORBEFWGTGYGKDIDWKGMLCALOE 1019
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

Db 1020 A---GGKLTLETYNSVNTFNGHITGRTKINFASRPSFLRWMTWGDQFCRERITQLOI 1076

```
Db      :: | : | | : | | : | | : | | : |  
1077. LKEKGM-VYQYNBKGDKDKEKETACIACYIKEMLTNWQDNKKGNQRV-TEVKGTSP-Y 113  
  
364 NSNGCDACTAY 304
```

```

Db      1134  KEDSDVKESKY 1144
          |  |  |

```

RESULT 6
8399
antigen-specific surface protein 1 homolog 3p7var1 - malaria parasite (Plasmodium falciparum) (Plasmodium falciparum)

C:\Database\15-Oct-1999 #sequence_level1510 15-Oct-1999
C:\Accession: T18399
R. Rubio, J. P.; Thompson, J. K.; Cowman, A. F.
EMBO J. 15, 4069-4077, 1996
#textChange 03-Jun-2000

A:Reference number: Z18927; MUID:96324414; PMID:8670911
A:Accession: J18399
A:Status: preliminary; translated from GH/EMBL/DBJ

A:Residues: 1-2042 <RUB>
A:Cross-references: EMBL:U05324; NID:g1297090; PID:g1297091;
C:Genetics:
C:Protein: 13

| | Score | DB 2: | Length |
|------------------------|------------|--------|--------|
| 23.28: | 395: | | 2042: |
| 26.28: | | | |
| Best local similarity: | pred No. 2 | 5e-20: | |

QY 24 KNIDNHSG-----ACMPPRKQLCVADLTGGGELRKRPEDILITFINCAKETHFAMHKY 78
 Db 746 RSADGAPSGDKGAICIPRRRRLLYLKLI-EGVDTTDDKSLRKWFIESAAVEFFFLMDRY 804

```

0Y 79 KK-----DNVNAENELKSGKIPGGRKOMYTGEGFRIEF 114
Db 805 KKLITPOSGSPLLCGMITGVGVENGEDDENNEKILLQGTETDGLFRMYTLGRDILE 864
0Y 115 GTD-----ISSCRVYKIDTSQITKSLGD--QATTEKGDIHIDN--KKLOEWMTI 160
Db 865 SGDDKKKNGSYDIIYSGDNVIREKRENTIKETIASEFFQNGKNEGTHYKKNVQFQPTWMD 924
0Y 161 HGPKIWEGMICALT-----NGISESEKK-----NIL 186
Db 925 NAKIHMHGMICALTYEEKTSSGDEKIEKDDAVYKKFEFGTPNGCNPLPQPTNGTSNBPI 984
0Y 187 QDYSYNKL-----NN-----AKKDCCCLEKPAKSGQFLRWYVENSDECRE 227
Db 985 SQYDYDOVVLKEENNANSTSPKSTSAAPSNDPFTTLQFLRPTERYERLEEWQONCRE 1044
0Y 228 RKKLEDKVEDYC-----IKANDYEGCKN-----NKSNN-----SCVVCKEEYVITGKK 272
Db 1045 RKKRLAQIYEDCGRGNDKYVSGDGECEVYRKQDYKSKISNFNCPCGCGRECRKRWMTIKTR 1104
0Y 273 TQYESQ-----EGKENTERKORKKPEYNSY-----KDASEY 304
Db 1105 TERPEQKKAYVDRKDKYTENKKGAEKSHHSISDQNFYKKLCTDY 1148

```

RESULT 7
T14029
variant-specific surface protein 1 - malaria parasite (plasmodium falciparum)

Nucleotide names: eynfncp
C:Species: Plasmodium falciparum
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text-change 09-Jun-2000
C:Accession: T14029

U. Exp. Med. 187, 15-23, 1998

A:Title: Identification of Plasmidum erythrocye membrane protein 1
 A:Reference number: 217860; MUID:98080592; PMID:9419207

A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-2228 <CHE> 20003473 20001457 20001450 20001450 20001450

| Query Match | Score | DB ID | Length |
|--------------------------------------|-------|----------|--------|
| C:Genetics: A:Note: FCR3S1.2-Var1 | 32.69 | SC000386 | 100000 |

| | | |
|----|---|--------|
| 20 | WDC-----KKINDSHSGA-CMPBRROKLCVRDL-----VGGGELRPK | 55 |
| | Best Local Similarity | 24.5%; |
| | Pred. NO. 1.76-19; | |
| | Matches 104; Conservative 59; Mismatches 123; | |
| | Indels 138; Gaps 14; | |

Db 917 WRCTPSEPTTSSDKNGAICVPPRRRRILYIKKIIVDMATKTESPOASGSEASSTSGSTTP 976

OY 56 P---EDILTRFINCAAKETHFAWHRKYKD-----NVAENELIKS 91

Db 977 PDSEKALLKAFVESAIETFFLLMHRKYEKKKAAVQEGAGCHGLPREVEESSPEYEDKLIKE 1036

Qy 92 GKLTPEGRKQMYTYFGDFPDIFPGTDISSCRYIKDTSQTIKSLGDAQTEKGDTHIDNN 151

Db 1037 GKIPDGLINQMFTLLGDYNDLEPSGSDNTTSVSKDPTSSSDNMLNIVLASSSTE-QER 1095
QY 152 KKL-----QEWWTTHGPKRIEHCALTF----- 174

Db 1096 EKMKYKKEIKNFRKCTERSAPNVLVSHQPTWMENNKGKTYIHGMVCAITLSDKIAKGVAKK 1155

175 -----NGLSESEKKNLQDYSYKNL-----NAEKDDCCLEKFAKK 210

QY

| | | | | |
|----|------|-------------------------------|--------------------------------------|------|
| DB | 1135 | P0K1ENPENLWDEANKKPKPPOYOYNTNK | LDENSNGTSPPRTQOASSDNTPTTLTHFVVR | 1215 |
| OY | 211 | P0E1RWYVMSDEFCRERK | -----KLEDKVE--DVC1KAKDYEKKNKNSN----- | 255 |

256 ---SCVVKACEYENYITGRKTOYESQEGKFNTERKQ-KPEYNSYSKR-----DAS 302

of EM

Db 1276 NCECCGHCSSYKRWIKRKEIEFKHOSNAYGOQKTDATRNNGTFFDKFCKILETWPDAA 1335
 QY 303 EYIK 306
 Db 1336 KFELE 1339

RESULT 8

T28157
 erythrocyte membrane protein 1 - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T28157
 R:Su: X.Z.; Heestwale, V.M.; Wertheimer, S.P.; Guinet, F.; Herfeldt, J.A.; Peterson, D.S.
 A:Title: The large diverse gene family var encodes proteins involved in cytoadherence at
 A:Reference number: Z20487; MUID:95330813; PMID:7606788
 A:Accession: T28157
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Accession: T28157
 A:Molecule type: DNA
 A:Residues: 12212 <YAN>
 A:Cross-references: EMBL:U031083; NID:91517813; PID:91517814; PIDN:AAB06961.1
 A:Genetics:
 A:Gene: EMP1

Query Match 22.5%; Score 383; DB 2; Length 2212;
 Best Local Similarity 28.5%; Pred. No. 2e-19;
 Matches 109; Conservative 44; Mismatches 102; Indels 128; Gaps 14;

QY 29 SHSGA-CMPRRQKLCVRDL-----TOGGEI-----RKPEDIL-TRKFIN 65
 Db 1337 SNSGSLICIPRRRLTYOKLHDWASGNTVYSGOAGTPQGTSPPSGKETPSDKLRTAFIQ 1396
 QY 66 CAAKETHFAWHKTKKDNVNAENELK-----SGKI 94
 Db 1397 SAAIEFFFLMDRYKKKEKEIEKKKAVANGLVSLNGCPPOPGVTGDSPOSKLOQTGYI 1456
 QY 95 PEGFRKQMYTFGDFRDIPEG--TDISSCRYIKDTSOTISKLGDOATTEK----- 143
 Db 1457 PPPFLROMYTLGADYADIFPGKNDI-----VIDTKNGDKDIAERKKIRKAIER 1505
 QY 144 ---GDTHTIDNKKILOEWMTIHGPKIWEGMLCALT-----NGI-----SES 180
 Db 1506 VLKNDSDOPSEKQGTWMEONGEHIWNCIMCALTYKKEDEKGTPLKONEGLKSLAMDEK 1565
 QY 181 EKKNTLODYSTKNNNAEKDDCC-----LEKFAKRPQFLRWYVMSDEFCRERK 229
 Db 1566 NKKRPDKQYQDKVKLDENSGTSPKTNHVPPTPLTNFISRPYRILEMEGETFCRERK 1625
 QY 230 KLEDEVEDYCI-----KAK---DYEGCKNNKSNM-----SCVAVCKEYENYITGK 271
 Db 1626 KRLEKIKVECMDEGCKKQCSGDGCEIRKQDYSTVDFYCPBCGYCRFYKRWIKK 1685
 QY 272 KTYVESQEGKNTERRKQKPEYN 294
 Db 1686 KDEYDKQKEAVNNKTDARNNNN 1708

RESULT 9

T28634
 variant-specific surface protein 7 - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T28634
 R:Su: X.Z.; Heestwale, V.M.; Wertheimer, S.P.; Guinet, F.; Herfeldt, J.A.; Peterson, D.S.
 A:Title: The large diverse gene family var encodes proteins involved in cytoadherence at
 A:Reference number: Z20487; MUID:95330813; PMID:7606788
 A:Accession: T28634
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-2182 <SOX>

A:Cross-references: EMBL:L42636; NID:9886379; PID:9886380; PIDN:AAA73399.1
 A:Genetics:
 A:Note: var-7

Query Match 21.5%; Score 367.5; DB 2; Length 2182;
 Best Local Similarity 25.4%; Pred. No. 2.6e-18;
 Matches 107; Conservative 60; Mismatches 128; Indels 127; Gaps 16;

QY 10 GCNPKRESYDWDG--KKINDNSHSG---ACMPRRQKLCVRDLTQ---GGE----- 52
 Db 874 GEGCKEKFPMKCVTPSGVSTATSGDGAICVPPRRRLRYVGLSQMASRGDETFTEVSS 933
 QY 53 -----IRKPEDILTRKFINCAKETHFAMHRYKKD-----NV 83
 Db 934 EATSAPOSSESEKRLATFESAIETFFLMHRYKKEKRPATODGAGLVSLEPSPGE 993
 QY 84 MNEEL-SSGKIPPEGFRKQMYTFSDFRDIF--GTDISSCRYIKDTSOTISKLGDOATT 141
 Db 994 DPOTLOQTGYIPDPFLROMFYTLADYKDIIYSGNSDTSDTGKQTPSSNDNLKNIVLE 1053
 QY 142 EKGDTHIDNK-----KIQEWMTIHGPKIWEGMLCA 172
 Db 1054 ASGSTDEQEKKKQIQAKIKITLNGATSGVPPVTKNSVKTPOQTWENIARDINAWACA 1113
 QY 173 LT-----NGLS-----ESEKNTLODYSYNKL-----NNAEKDDC----- 202
 Db 1114 LTYKENDARGTSAKIEQNKDKLALMDKANNTPIEKQYTNVNLKEDSGAKSNDTIOP 1173
 QY 203 CLEKASRPQFLRWYVMSDEFCRERKLEKVEDVCIK---ANDYEG---CKNKS-- 253
 Db 1174 TLKNFVEIPTFERMLHEWNSFCFERAKRLAQIKHECMDEDEGKQYSGDGEYCEIFSKQ 1233
 QY 254 -----NNSCVYKVEYENYITGKKTYVESQEGKFNKK-----RQKPEYNSKSDAS 302
 Db 1234 YNVLDLSSSCAKRRLYKTIETKKTYERQQAQNAEQKSNYENEDKQCTQTSNNAN 1293
 QY 303 EY 304
 Db 1294 EF 1295

RESULT 10

T28161
 hypothetical protein FC93-varT11-1 - malaria parasite (Plasmodium falciparum) (fragme
 C:Species: Plasmodium falciparum
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T28161
 R:Hernandez-Rivas, R.; Mattei, D.; Sterkers, Y.; Peterson, D.S.; Wellens, T.E.; Scher
 M.Ol. Cell. Biol. 17, 604-611, 1997
 A:Title: Expressed var genes are found in Plasmodium falciparum subtelomeric regions.
 A:Reference number: Z20483; MUID:97154495; PMID:9001213
 A:Accession: T28161
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-2647 <HER>
 A:Cross-references: EMBL:U067959; NID:91794255; PID:91809295; PIDN:AAC47438.1
 A:Experimental source: strain FCQ27/PNG
 A:Genetics:
 A:Introns: 2158/3
 A:Note: FC93-varT11-1

Query Match 21.2%; Score 361.5; DB 2; Length 2647;
 Best Local Similarity 27.4%; Pred. No. 8.7e-18;
 Matches 107; Conservative 44; Mismatches 104; Indels 135; Gaps 14;

QY 29 SHSGA-CMPRRQKLCVRDL-----TOGGEI-----RKPEDIL-TRKFIN 65
 Db 1338 SNSGSLICIPRRRLTYOKLHDWASGNTVYSGOAGTPQGTSPPSGKETPSDKLRTAFIQ 1397
 QY 66 CAAKETHFAWHKTKKDNVNAENELKSGK-----I 94
 Db 1398 SAAIEFFFLMDRYKKKGAIAKKKKQMDYSLSTADPHNPNVSLVIAPNPNYKTKCVI 1457

A:Molecule type: DNA
 A:Residues: 1-2197 <GAR>
 A:Cross-references: GB:AE001434; GB:AE001362; NID:93845341; PIDN:AAC71996.1; PID:9384534
 A:Experimental source: clone 307
 C:Genetics:
 A:Gene: PFB1055C

Query Match 20.1%; Score 342.5; DB 2; Length 2197;
 Best Local Similarity 24.6%; Pred. No. 1.7e-16;
 Matches 109; Conservative 62; Mismatches 121; Indels 151; Gaps 16;

10 GCNPKESYPMDCCKKINDNS-----HSGA-----CMPPRKQKLCVRLD-----48
 936 GPGGKEREPNWKCVSSGKSVATAGSSGATGSGDKGALCVPPRRRLVYGLTKLTSAG 995
 49 -----GG-----EIRKPEDILTFINCAKETHPAMRYKK-----80
 996 TSESPGSGSESSRASVSGNGDDITTTESLRKWFETLTALETFFLMHRYKWEAOK 1055
 81 -----DNVNAENEL-KSGKIPGFRKQMYTFGDFRDIFF-----114
 1056 KAEIQRNGLLGTGASINLGDDSNPOTQLOKSGTIPDLPLRLMFTLGDYRDIYRGVA 1115
 115 -----GTDI-----SSGRYIKDTISQTKSLGDOATTEKGD-----THIDNKKLQEMW 158
 1116 DDKNGGNITLNASGNKDEKQKMEKIOEKI-EQILPTSGNKETRGFQNSYNDROSL--W 1171
 159 TTHGPKIMEGMLCALT-----NGL-----SESEKK-----NLL 186
 1172 DRAIEHVHGMVLCALTYKDDNGLKGVKKPKQIENPEKLMNNTTKKPKYDEKYQYOTAKL 1231
 187 QDYSYKLNNAEKDDCCLEKFAKPOFLRMVYEWSEDFCRERKKLEKVEDVCIR-----241
 1232 EDESGERKPDSSASGTKLDFIKRPYFYLEEMGECNFKKRTKEMGLIKEOCYKNGGRC 1291
 242 -----AKDYEGCKNNKSNNSGVCKYEYENITTKKQYQESQEGKFNTERKOKRP 291
 1292 SGDGKLKNEIVIDKEIFEDLCLPTCARHCRTYKWKINTKRDENFKNSAAYEQKKYEE 1351
 292 EYNSYSK-----DASEYL 305
 1352 ENDSAKKNGVCGTLKDAAEFL 1374

RESULT 14

T18378

variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)

A:Alternate names: erythrocyte membrane binding protein 1 (EMP1)

A:Species: Plasmodium falciparum

A:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

A:Accession: T18378

R:Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschi, T.

Cell 82, 77-87, 1995

A:Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and

A:Reference number: Z18925; MUID:95330812; PMID:7541722

A:Accession: T18378

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2924 <BAR>

A:Cross-references: EMBL:U27338; NID:9914918; PID:9914919; PIDN:AAB60251.1

A:Genetics: EMP1

A:Gene: EMP1

A:Introns: 2476/3

Query Match 18.2%; Score 311; DB 2; Length 2924;

Best Local Similarity 23.0%; Pred. No. 4.3e-14;

Matches 105; Conservative 60; Mismatches 120; Indels 172; Gaps 18;

14 KESYPMDCCKKINDNS-----GA-----CMPPRKQKLCVRLD-----47
 860 KTAFTSMKCISSGNNITTESTTKFOAGCTPGSKDTGSCVPPRRRKLTVGLHDMAGGET 919
 48 -----TQGE-----IRKPE-DILTFINCAKETHPAMW 76

DB 920 TEAKSOETSGGQTPSGNESSPSEKLPQGPTEETKETPESSILHAFVSPRRRLRFLPW 979
 QY 77 KYRKD-----NVNAENELKSGKIPGFRKQMYTFGDFRDIFF 113
 DB 980 KFEQWKAQHAGATGQOTIIGTLDGGGEETEDKLTHTHIPPDLRQMFYTLGSDYRDL 1039
 QY 114 FG-TDI-----SSGRYIKDTISQTKSLGDOATTEKGD-----HIDNKKLQEMW 158
 DB 1040 VGNITDLYVHTSGNKEDMOIMEAIOKKI-EQILPTSGSSPSRPYOTQSHVENPR-KTWW 1097
 QY 159 TTHGPKIMEGMLCALTNG-----SESKNNIIDYSYNKLNNAEK 199
 1098 NENKKGIMGVCALTYNTDTPSGTAPQIOEVRKIKDENSEKNPKIQYKYDYKIDDT 1157
 200 DDC-----CLEKFAKPOFLRMVYEWSEDFCRERKKLEKVEDVCIRAK-D 244
 DB 1158 SDAKTGSPVPSGKLTPLTDFISRPYFYLEEMGECNFKKRTKEMGLIKEOCYKNGGRC 1217
 QY 245 YEGC-----KNK-----SNNSGVCKYEYENITTKKQYQESQEGKFNTERK 287
 DB 1218 HEHCSGDGYDCTRTDADRNDKRVDLNCRDCHIOCKRYKRWIDIKFDEYHKOEKY-----1272
 QY 288 OKKPEYNSYSK-----KDASEYLKD 307
 DB 1273 --QGEYDKLTDKDSSGDNCKDIEKHKSAVFLKE 1307

RESULT 15

T18396

erythrocyte membrane protein variant 2 - malaria parasite (Plasmodium falciparum) (fr

A:Species: Plasmodium falciparum

A:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

A:Accession: T18396

R:Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschi,

Cell 82, 77-87, 1995

A:Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen a

A:Reference number: Z18925; MUID:95330812; PMID:7541722

A:Accession: T18396

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1729 <BAR>

A:Cross-references: EMBL:U27339; NID:9914920; PID:9914921; PIDN:AAB9134.1

A:Genetics: EMP1

A:Note: var-2

Query Match 17.5%; Score 298.5; DB 2; Length 1729;
 Best Local Similarity 23.1%; Pred. No. 1.9e-13;
 Matches 107; Conservative 55; Mismatches 126; Indels 175; Gaps 20;

14 KESYPMDCCKKINDNS-----DNHSGA-CMPPRKQKLCVRLD-----T 48
 DB 860 KTAFTSMKCISSGNNITTESTTKFOAGCTPGSKDTGSCVPPRRRKLTVGLHDMAGGETTEAKSOET 919
 QY 49 QGGE-----IRKPE-DILTFINCAKETHPAMRYKK-----80
 DB 920 SCGQKTPSGNESSPSEKLPQGPTEETKETPESSILHAFVSPRRRLRFLPWKFKQWKA 979
 QY 81 -----DNVNAENELKSGKIPGFRKQMYTFGDFRDIFFGTDISSGC 122
 DB 980 QHGAGATGLOLPGVTVDDSDPDPOTLKRGINBNDFLRQMFYTLGDRDICTGCD--R 1035
 123 YI-----KQTSQIKSKLGDQATTEK-----DTHIDNKKLQ 155
 DB 1036 DIVGDTIVSTTEBESIKKTSKISITTEGLKQYVTSPPRDTSSRPVHROTQSVETPQ-Q 1094
 QY 156 EWTTHGPKIMEGMLCALT-----NGLSESEKK-----NIDYSYNKLNNAE--KDDCC 203
 DB 1095 TWMEANGPHIMNMCALTYEDSGAIGQPPQYKEDADVLEKIKRPTANGIKWYLKEDWT 1154
 QY 204 -----LEKFAKPOFLRMVYEWSEDFCRER-KLEKVEDVCIT 240

Db 1155 SSAMPTSSSSSGSNDPINTFPKLTJFVFIPTFFRYLHEWGONFCERMRRLKQIYKECKV 1214
OY 241 KADY-----EGCKNNKSNN-----CYKVCKEYENYITGKKTO 274
Db 1215 GENGIGRGKOKTPQCCSCYGEDDQLSKYSYDTVADIECPKCAKHCRWYRWIEKKDE 1274
OY 275 YESOEGKFNTEK-----ROKKPEYNSY-----SKKDASEYIK 306
Db 1275 FTBQEKAFPRKOKDYVNGNNGKGGDNGFCITLKSISDAQPLE 1317

Search completed: June 20, 2003, 15:06:23
Job time : 26.3035 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 14:37:29 ; Search time 13.5539 Seconds

(without alignments)
942.514 Million cell updates/sec

Title: US-10-087-013-11

Perfect score: 1706

Sequence: 1 GNGSNEISGNCNPRESEYDWN.....KKPEYNSYKRDASEYLKDK 308

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

a1 number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt.40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 271.5 | 15.9 | 1070 | 1 | PVDG_PLAKN |
| 2 | 239 | 15.2 | 1070 | 1 | PVDR_PLAYS |
| 3 | 251 | 14.7 | 1435 | 1 | EBAL_PLAC |
| 4 | 246.5 | 14.4 | 1153 | 1 | PVDB_PLAKN |
| 5 | 227.5 | 13.3 | 1073 | 1 | PVDA_PLAKN |
| 6 | 109 | 6.4 | 1118 | 1 | UBP8_HUMAN |
| 7 | 104 | 6.1 | 2492 | 1 | ATRX_HUMAN |
| 8 | 100 | 5.9 | 278 | 1 | YD33_SCHPO |
| 9 | 99.5 | 5.8 | 941 | 1 | YD33_MOUSE |
| 10 | 99 | 5.7 | 1181 | 1 | YD33_MOUSE |
| 11 | 97.5 | 5.7 | 944 | 1 | YD33_MOUSE |
| 12 | 97.5 | 5.7 | 1170 | 1 | SMC2_YEAST |
| 13 | 97.5 | 5.7 | 1407 | 1 | CTAA_DICDI |
| 14 | 97 | 5.7 | 392 | 1 | DLTD_BACSU |
| 15 | 96.5 | 5.7 | 348 | 1 | CYL2_HUMAN |
| 16 | 96.5 | 5.7 | 959 | 1 | VDP_RAT |
| 17 | 96.5 | 5.6 | 507 | 1 | SPT1_MOUSE |
| 18 | 96 | 5.6 | 750 | 1 | YD33_MOUSE |
| 19 | 96 | 5.6 | 1162 | 1 | BXEN_CLOBU |
| 20 | 96 | 5.6 | 1391 | 1 | MST2_DROH |
| 21 | 96 | 5.6 | 1875 | 1 | MLP1_YEAST |
| 22 | 96 | 5.6 | 2104 | 1 | MYX3_SCHPO |
| 23 | 96 | 5.6 | 2195 | 1 | SC16_YEAST |
| 24 | 95.5 | 5.6 | 1484 | 1 | CES2_HUMAN |
| 25 | 95 | 5.6 | 1341 | 1 | RPA1_METUA |
| 26 | 94.5 | 5.5 | 526 | 1 | UBPW_MOUSE |
| 27 | 94.5 | 5.5 | 650 | 1 | BM86_BOOMI |
| 28 | 94.5 | 5.5 | 743 | 1 | ABRA_PLAC |
| 29 | 94 | 5.5 | 415 | 1 | TRMU_SCHPO |
| 30 | 94 | 5.5 | 452 | 1 | TM44_HUMAN |
| 31 | 93.5 | 5.5 | 607 | 1 | HTPG_FUSNN |
| 32 | 93.5 | 5.5 | 804 | 1 | ME46_HUMAN |
| 33 | 93.5 | 5.5 | 962 | 1 | VDP_HUMAN |

| | | | | | | |
|----|------|-----|------|---|------------|--------------------|
| 34 | 93.5 | 5.5 | 1095 | 1 | NEB1_RAT | O35867 ratus norv |
| 35 | 93.5 | 5.5 | 2869 | 1 | RBP1_PLAVB | O00798 plasmodium |
| 36 | 93 | 5.5 | 272 | 1 | Y603_AOUAE | O66861 aquilex aeo |
| 37 | 92.5 | 5.4 | 555 | 1 | EDC8_DAICA | P20075 dancus caro |
| 38 | 92.5 | 5.4 | 742 | 1 | NEB1_HUMAN | O901j8 homo sapien |
| 39 | 92.5 | 5.4 | 824 | 1 | AD17_HUMAN | P78536 homo sapien |
| 40 | 92 | 5.4 | 853 | 1 | YCG1_YEAST | P25588 saccharomyc |
| 41 | 92 | 5.4 | 902 | 1 | MOPT_ARATH | P93024 arabidopsis |
| 42 | 92 | 5.4 | 2476 | 1 | ATRX_MOUSE | O61687 mus musculu |
| 43 | 92 | 5.3 | 1162 | 1 | BXEN_CLOBU | P46082 clostridium |
| 44 | 91 | 5.3 | 1230 | 1 | SMC3_YEAST | P47037 saccharomyc |
| 45 | 91 | 5.3 | 1584 | 1 | YU9G_YEAST | P47170 saccharomyc |

ALIGNMENTS

| RESULT 1 | ID | PVDG_PLAKN | STANDARD: | PRT: | 1070 AA. |
|----------|--|------------|--------------------|-------------------------------------|---------------------|
| AC | P50494: | | | | |
| DT | 01-OCT-1996 (Rel. 34, Created) | | | | |
| DT | 01-OCT-1996 (Rel. 34, Last sequence update) | | | | |
| DT | 01-OCT-1996 (Rel. 34, Last annotation update) | | | | |
| DE | Duffy receptor, gamma form precursor (Erythrocyte binding protein). | | | | |
| OS | Plasmodium knowlesi. | | | | |
| OC | Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium. | | | | |
| OX | NCBI_TaxID=5650; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RA | MEDLINE=92357776; PubMed=1496004; | | | | |
| RA | Adams J.H., Slim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.; | | | | |
| RT | "A family of erythrocyte binding proteins of malaria parasites."; | | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089(1992). | | | | |
| CC | - FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP DETERMINANT. | | | | |
| CC | - SUBCELLULAR LOCATION: Type I membrane protein. | | | | |
| CC | - SIMILARITY: HIGH TO P.VIVAX DUFFY RECEPTOR. | | | | |
| CC | - This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch). | | | | |
| CC | EMBL; M90695; AAA29604.1; - | | | | |
| DR | Malaria; Receptor; Glycoprotein; Signal; Transmembrane; | | | | |
| KW | Multigene family. | | | | |
| FT | SIGNAL | 1 | 21 | POTENTIAL. | |
| FT | CHAIN | 22 | 1070 | DUFFY RECEPTOR, GAMMA FORM. | |
| FT | DOMAIN | 22 | 1003 | EXTRACELLULAR (POTENTIAL). | |
| FT | TRANSMEM | 1004 | 1025 | POTENTIAL. | |
| FT | DOMAIN | 1026 | 1070 | CYTOPLASMIC (POTENTIAL). | |
| FT | CARBOHYD | 134 | 134 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | CARBOHYD | 179 | 179 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | CARBOHYD | 676 | 676 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | CARBOHYD | 743 | 743 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | CARBOHYD | 785 | 785 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | CARBOHYD | 936 | 936 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | SITE | 279 | 281 | CELL ATTACHMENT SITE (POTENTIAL). | |
| SO | SEQUENCE | 1070 AA; | 120931 MW; | 70306881B011B50 CRC04; | |
| QY | Query Match | 15.9%; | Score 271.5; | DB 1; | Length 1070; |
| QY | Best Local Similarity | 27.5%; | Pred. No. 4.2e-12; | | |
| DB | Matches | 90; | Conservative 39; | Mismatches 117; | Indels 81; Gaps 16; |
| QY | 5 SNEISGNCNP--KESIPDMDCKKNIDNSHGACMPRRKRLCVRLDTGGELIRKPEDILTK 62 | | | | |
| DB | 208 NNVMDCCKDKRRRGGRDWCCTPERD--VCIPRRVQLDIMEIT----- 248 | | | | |
| QY | 63 FINCAAKETHPAMHK-----YKRD-----NVNAENELKSGK---IPGFRKQMYVTFG 107 | | | | |

```

Db 249 --NLVPTDHF--HSDIIRKSTSRRLIYDVGREDLLKTKYNNVSEDLCKDIMSIO 304
QY 108 DFDIDFEGTDISSGRYIKDTSQIKSKLGDQATTEKGDTHIDNKKLOEWMTIHGPKIME 167
Db 305 DFDIDIMGTDMGIGSLVQNNLRSLFPGTISAL-----DRKK--WMDHKKDINK 355
QY 168 GMLCALTNGLSSEKKNIIQDYSYKNLNNAEKDCCLEKFAK--DQFLRWYVMS---- 221
Db 356 AM-----LSYKKNR-----YSAMN-----CKEDVQIVKEPOIYRWIREGCRDY 396
QY 222 DEFCEERKKLEDEVEVCIAKDYEGCKNNKSNNSGVKCKEVENITGKTOYESQEK 281
Db 397 SERERKRKLNECED-----KLYSTMLICTLPPCNACKSYDEMITSKKQWDLSTR 451
QY 282 FTEKRQKRPENYSYSKDASEYLKDK 308
Db 452 FSSVKAQKILETENIAR--AYDILKE 476

```

RESULT 2

```

DR PLAYS STANDARD: PRT: 1070 AA.
ID PVDRL PLAYS
AC P22290;
DT 01-FEB-1991 (Rel. 19, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Duffy receptor precursor (Erythrocyte binding protein).
GN PVDRL
OS Plasmodium vivax (strain Salvador I).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=126793;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91187056; PubMed=1849231;
RA Fang X., Kaslow D.C., Adams J.H., Miller L.H.;
RT "Cloning of the Plasmodium vivax Duffy receptor."
RL Mol. Biochem. Parasitol. 44:125-132(1991).
CC - FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP DETERMINANT.
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - SIMILARITY: HIGH, TO P. KNOWLESI DUFFY RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```

```

DR EMBL: M61095; AAA63423.1;
KM Malaria; Receptor; Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 20
FT CHAIN 21 1070
FT TRANSMEM 21 1007
FT DOMAIN 1008 1025
FT DOMAIN 1026 1070
FT CARBOHYD 183 183
FT CARBOHYD 255 255
FT CARBOHYD 351 351
FT CARBOHYD 420 420
FT CARBOHYD 715 715
FT CARBOHYD 787 787
FT CARBOHYD 825 825
FT CARBOHYD 903 903
FT CARBOHYD 938 938
SQ SEQUENCE 1070 AA; 119683 MW; CB051DPI3E294603 CRC64;

```

Query Match 15.2%; Score 259; DB 1; Length 1070;
 Best Local Similarity 24.9%; Pred. No. 3.3e-11;
 Matches 89; Conservative 38; Mismatches 127; Indels 104; Gaps 14;

```

QY 1 GNDGSENLSCNP-----KESYPMDCCKNNDS 29
Db 178 GNSRKNSSNGNGNDYDIDHKKTISATINAFLONTVMKNCNKKRRRREDWMC-----NF 232
QY 30 HSGACPPPRQYLCYRDLT---QGEIRKPEDL-----TKFNCAKETHFAW---- 75
Db 233 KRDVCIPDRRYOLCMKELTNLVNNDTNHHRIDTFERKLYLKRKLLIYDAVFEGLLLKNN 292
QY 76 HKYKKNVNAENELSGKIPESGRKQMYTFEDFDIDFEGDIDISSCRYIKDTSQIKSKL 135
Db 293 YRYNMD-----FCKDIRWSLGDIEDDIDIMGTDMGIGSLVQNNLRSLF 336
QY 136 GDOATTEKGDTHIDNKKLOEWMTIHGPKIMEGMLCALTNGLSSEKKNIIQDYSYKNLN 195
Db 337 G---TDEKA-----QQRKQWNNESKQAIWTAMYSYKRL-----KGNFTICLNVAV 383
QY 196 NAEKDDCCLEKFAKPOFLRWYVMSDEFCERKKLEKVEDVC---TKAKDYEGCKNNK 252
Db 384 NIE-----POIYRWIREGCRDYSELPTREVOKRKCKGKINTDKKVC--- 428
QY 253 SNNSGVKCKEVENITGKTOYESQEGF---NTEKRQKRPENYS--KDASEY 304
Db 429 -VPCQNAKCKSYDQWITRKKNQWDLNKNFISVKNNAEKVQTAGIVTPYDILKEIDEF 485

```

RESULT 3

```

EBAL PLAF STANDARD: PRT: 1435 AA.
ID EBAL PLAF
AC P19214;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Erythrocyte-binding antigen EBA-175.
OS Plasmodium falciparum (isolate Camp / Malaysia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5835;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=90377299; PubMed=2204835;
RA Sim B.K.L.;
RT "Sequence conservation of a functional domain of erythrocyte binding antigen 175 in Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 41:293-296(1990).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL: X52524; CA36756.1;
DR PIR: S11561; S11561.
KW Antigen.
FT DOMAIN 159 1104
FT VARIANT 1031 1031
FT SEQUENCE 1435 AA; 167389 MW; 32A4309021BIC3D6 CRC64;

```

```

DR EMBL: X52524; CA36756.1;
DR PIR: S11561; S11561.
KW Antigen.
FT DOMAIN 159 1104
FT VARIANT 1031 1031
FT SEQUENCE 1435 AA; 167389 MW; 32A4309021BIC3D6 CRC64;

```

Query Match 14.7%; Score 251; DB 1; Length 1435;
 Best Local Similarity 26.3%; Pred. No. 1.7e-10;
 Matches 79; Conservative 47; Mismatches 100; Indels 74; Gaps 14;
 QY 20 WDCKKNIDNSISGACMPPRQKLCYRDLTGGGIRKPED-----ILTKFNCAKETHFA 74
 Db 474 WECKNPYTLSTKDYCVPRRQELC-----GNIDRIYDNLMIKELHIAIAYESRIL 527
 QY 75 WHRYKKNVNAENELSGKIPESGRKQMYTFEGDFRIFPGDIDISSCRYIKDTSQIKSK 134
 Db 528 KRRK-----NKDREYCKTIINKTFDIRIDIGTD-----YNNDLN--RKL 568
 QY 135 LGDQATTEKGDTHIDNKK-----LOEWMTIHGPKIMEGMLCALTNGLSSEKKNIIQDY 189

FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 348 348 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 746 746 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 779 779 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 788 788 N-LINKED (GLCNAC. . .) (POTENTIAL)
 SQ SEQUENCE 1073 AA; 120683 MW; 3965FC9F46B71808 CRC64;

Query Match 13.3%; Score 227.5; DB 1; Length 1073;
 Best Local Similarity 25.5%; Pred. No. 66-09; Indels 63; Gaps 12;
 Matches 77; Conservative 38; Mismatches 124;

QY 5 SNEISGCP- -KESYPMDCCKNNINSHGACMPPRROKLCVRYDLT---QGGEIRKPPDI 59
 DB 208 NNVMDCKNDKRKRGERDMDCPAERD-----ICISVRRQCKGLTLNVTNRITSHNDI 262
 QY 60 LTKFNCNAKETHFAMHK-----YKDNVNAENELKSGKIPESGRKOMYTFGDFPDIFFG 115
 DB 263 TFLKLNLRKRLMYDAVAGDILLKKNYQYNKE-----FCKDIRMGGLGDFGDIIMG 313
 QY 116 TDISSCRYIKDTSQITKSKLGDQATEKGDTIIDNKKLOEWMTIHGKIMEGMLCALTN 175
 DB 314 TMMEIGISQYVENNLROYFG---TDERAK---QDRK---QWNESKEHWRAMMFSTIRS 364
 QY 176 GLSESEKKNILQDYSYKLNNAEKDCCLEKFAKSPQFLRMVVEWSEDFECERKKLEDKV 235
 DB 365 RLKE-----KFWICKKDYVLK---VEPQYRRIRIREGRIMYSKLEKREGOKL 408
 QY 236 EDVCIKADYEGCKNNKNSNS-----CVKCKEYENYITGKTQYESQEGKENTEKRO 288
 DB 409 NECKASKLYY-----NNMAICMLPLCHDACKSYDQMITRKKKQMDVLSTKSSVAKT 460
 QY 289 KK 290
 DB 461 QK 462

RESULT 6
 UBP8_HUMAN
 ID UBP8_HUMAN STANDARD; PRT; 1118 AA.
 AC P40818;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ubiquitin carboxyl-terminal hydrolase 8 (EC 3.1.2.15) (Ubiquitin
 thiolesterase 8) (Ubiquitin-specific processing protease 8)
 DE (Deubiquitinating enzyme 8).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RX MEDLINE=96051398; PubMed=7584044;
 RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
 RA Seki N., Kawarayashi Y., Ishikawa K.-I., Tabata S.,
 RT "Prediction of the coding sequences of unidentified human genes. II.
 RT The coding sequences of 40 new genes (K1AA0041-K1AA0080) deduced by
 RT analysis of cDNA clones from human cell line KG-1.";
 RL DNA Res. 1:223-229(1994).
 CC -i- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
 CC ubiquitin + a thiol.
 CC -i- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: D29956; BAA06225.1;
 CC DR MEROPS: C19.011;
 CC DR Genes: HGNC:12631; USP8.
 CC MIM: 603158;
 CC DR InterPro: IPR001763; Rhodanese-like.
 CC DR InterPro: IPR001394; UCH-2.
 CC DR Pfam: PF00442; UCH-1;
 CC DR Pfam: PF00443; UCH-2; 1.
 CC DR Pfam: PF00581; Rhodanese; 1.
 CC DR SMART: SM00450; RHOD; 1.
 CC DR PROSITE: PS00972; UCH_2_1; 1.
 CC DR PROSITE: PS00973; UCH_2_2; 1.
 CC DR PROSITE: PS0235; UCH_2_3; 1.
 CC DR Ubl conjugation pathway; Hydrolyase; Thiol protease; Multigene family.
 FT ACT_SITE 786 786 BY SIMILARITY.
 FT ACT_SITE 1059 1059 BY SIMILARITY.
 FT ACT_SITE 1067 1067 BY SIMILARITY.
 SQ SEQUENCE 1118 AA; 127523 MW; 8B884B7A842F9A9A CRC64;

Query Match 6.4%; Score 109; DB 1; Length 1118;
 Best Local Similarity 21.6%; Pred. No. 1.9;
 Matches 66; Conservative 34; Mismatches 121; Indels 84; Gaps 14;

QY 60 LTKFNCNAKETHFAMHK-----KDNVNAENEL-KSKTIFEGF-----R 99
 DB 788 MNSILOCLCNAPHLADFNRCYODDINRNLGHKCEVAEFGIIMKALMTQYRISP 847
 QY 100 KQMYTFGDFPDIFGDISSCRYIKDTSQITKSKLGD-----QATTEGDT 147
 DB 848 KDEITITGKINDQFAG-----YSQDSQELLFLMDGHEIDNKADNRKRYKEENVD 900
 QY 148 IDNKKLOEWMTIHGPIWEGMLCALTNGLSES-----EKNILQDYSYKLNNAEK 199
 DB 901 LDDFKAEHAWOKH-KOLNESIIYALFOGFKSTVCLTCHKRSRFEAFMYLSLPLAST 959
 QY 200 DDG-----CLEKFAKSPQFL---RWYVMSDFECERKKLEDKVD-----VCIKARD 244
 DB 960 SKCTLDQCLRLFSEKELTNNRNYC-----SHCRARDSLKTKIWKLPVLLVHKRRS 1015
 QY 245 YEGCKNNKNSNCVKKCKEYEN-----YINGKKTQYESQEGKENTEKROKKPEYNSYRK 299
 DB 1016 YDGRWKKQKLTYS---VDPLPENNLSQYVIGPRN-----NLKKYNLFSSVNHGCL 1063
 QY 300 DASRY 304
 DB 1064 DGGHY 1068

RESULT 7
 ATRX_HUMAN
 ID ATRX_HUMAN STANDARD; PRT; 2492 AA.
 AC P46100; P51068; Q15886; Q9NTS3; Q9H0Z1;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Transcriptional regulator ATRX (X-linked helicase II) (X-linked
 DE nuclear protein) (XNP) (Znf-HX).
 GN ATRX OR RAD54L OR XH2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5), VARIANT S-1860, AND
 RP VARIANTS ATR-X.
 RX MEDLINE=97123494; PubMed=8968741;
 RA Picketts D.J., Higgs D.R., Bachoo S., Blake D.J., Quarrell O.W.T.,
 RA Gibbons R.J.,
 RT "ATRX encodes a novel member of the SNF2 family of proteins: mutations
 RT point to a common mechanism underlying the ATR-X syndrome.";

Hum. Mol. Genet. 5:1899-1907(1996).
[12]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).
RX MEDLINE-97386582; PubMed-9244431;
RA Villard L., Lossi A.-M., Cardoso C., Proud V., Chiaroni P.,
RA Colleaux L., Schwartz C., Fontes M.;
RT "Determination of the genomic structure of the XNP/ATRX gene encoding
RT a potential zinc finger helicase.";
RL Genomics 43:149-155(1997).
[13]
RP SEQUENCE OF 860-2492 FROM N.A.
RX MEDLINE-95179111; PubMed-7874112;
RA Stayton C.L., Dabovic B., Gulsano M., Geetz J., Broccoli V.,
RA Giovannazzi S., Bossolasco M., Monaco L., Rastan S., Boncinelli E.,
RA Bianchi M.E., Consalez G.G.;
RT "Cloning and characterization of a new human Xq13 gene, encoding a
RT putative helicase.";
RL Hum. Mol. Genet. 3:1957-1964(1994).
[14]
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
MEDLINE-94214473; PubMed-8162050;
Geetz J., Pollard H., Consalez G., Villard L., Stayton C.L.,
RA Millasseau P., Khrestchatsky M., Fontes M.;
RT "Cloning and expression of the murine homologue of a putative human
RL X-linked nuclear protein gene closely linked to PK1 in Xq13.3.";
RL Hum. Mol. Genet. 3:39-44(1994).
[15]
RP SEQUENCE OF 2401-2492 FROM N.A. AND VARIANTS ATR-X.
RX MEDLINE-95211835; PubMed-7697714;
Gibbons R.J., Picketts D.J., Villard L., Higgs D.R.;
RT "Mutations in a putative global transcriptional regulator cause X-
RT linked mental retardation with alpha-thalassemia (ATR-X syndrome).";
RL Cell 80:837-845(1995).
[16]
RP SEQUENCE OF 1375-2492 FROM N.A.
RA Pearce A., Chapman J.;
RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
[17]
RP EZH2 BINDING.
MEDLINE-98167853; PubMed-9499421;
RA Cardoso C., Timsit S., Villard L., Khrestchatsky M., Fontes M.,
RA Colleaux L.;
RT "Specific interaction between the XNP/ATR-X gene product and the SET
RT domain of the human EZH2 protein.";
RL Hum. Mol. Genet. 7:679-684(1998).
[18]
RP SUBCELLULAR LOCATION, AND ASSOCIATION WITH PERICENTROMERIC
HETEROCHROMATIN.
MEDLINE-20040663; PubMed-10570185;
McDowell T.L., Gibbons R.J., Sutherland H., O'Rourke D.M.,
RA Bickmore W.A., Pombo A., Turley H., Gatter K., Picketts D.J.,
RA Buckle V.J., Chapman L., Rhodes D., Higgs D.R.;
RT "Localization of a putative transcriptional regulator (ATRX) at
RT pericentromeric heterochromatin and the short arms of acrocentric
RT chromosomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:13983-13988(1999).
[19]
RP DISEASE.
MEDLINE-20213147; PubMed-10751095;
RA Villard L., Fontes M., Ades L.C., Geetz J.;
RT "Identification of a mutation in the XNP/ATR-X gene in a family
RT reported as Smith-Rimman-Myers syndrome.";
RL Am. J. Med. Genet. 91:83-85(2000).
[10]
RP VARIANT ATR-X SFR-1713.
RX MEDLINE-97196774; PubMed-9043863;
RA Villard L., Lacombe D., Fontes M.;
RT "A point mutation in the XNP gene, associated with an ATR-X phenotype
RT without alpha-thalassemia.";
RL Eur. J. Hum. Genet. 4:316-320(1996).
[11]
RP VARIANT JM GIN-2131.
MEDLINE-96224392; PubMed-8630485;
RA Villard L., Geetz J., Mattei J.-F., Fontes M., Saugier-Verber P.,
RA Munnich A., Lyonnet S.;
RT "XNP mutation in a large family with Juberger-Marsidi syndrome.";
RL Nat. Genet. 12:359-360(1996).
[12]
RP VARIANTS ATR-X.
RX MEDLINE-97467722; PubMed-9326931;
RA Gibbons R.J., Bachoo S., Picketts D.J., Aftimos S., Asenbauer B.,
RA Bergoffen J., Berry S.A., Dahl N., Fryer A., Keppeler K., Kurosawa K.,
RA Levin M.L., Masuno M., Meri G., Pierpont M.E., Stanley S.F.,
RA Higgs D.R.;
RT "Mutations in transcriptional regulator ATRX establish the functional
RT significance of a PHD-like domain.";
RL Nat. Genet. 17:146-148(1997).
[13]
RP VARIANT ATR-X LEU-246.
RX MEDLINE-20123062; PubMed-10660327;
RA Flichera M., Romano C., Castiglia L., Falla P., Ruberto C., Amata S.,
RA Greco D., Cardoso C., Fontes M., Ragusa A.;
RT "New mutations in XNP/ATR-X gene: a further contribution to
RT genotype/phenotype relationship in ATR/X syndrome.";
RL Hum. Mutat. 12:214-214(1998).
[14]
RP VARIANT SHS LVS-1742.
RX MEDLINE-99347960; PubMed-10417298;
RA Lossi A.-M., Millan J.M., Villard L., Orellana C., Cardoso C.,
RA Prieto F., Fontes M., Martinez F.;
RT "Mutation of the XNP/ATR-X gene in a family with severe mental
RT retardation, spastic paraplegia and skewed pattern of X inactivation:
RT demonstration that the mutation is involved in the inactivation
RT bias.";
RL Am. J. Hum. Genet. 65:558-562(1999).
[15]
RP VARIANT CMS THR-2050.
RX MEDLINE-99326061; PubMed-10398237;
RA Abidi F., Schwartz C.E., Carpenter N.J., Villard L., Fontes M.,
RA Curtis M.;
RT "Carpenter-Waziri syndrome results from a mutation in XNP.";
RL Am. J. Med. Genet. 85:249-251(1999).
[16]
RP VARIANTS ATR-X E-175; 179-V-K-198 DEL; S-190; P-219; L-246 AND C-249.
RX MEDLINE-99219535; PubMed-10204841;
RA Villard L., Bonito M.-C., Abidi F., Ragusa A., Belongue J.,
RA Lossi A.-M., Seaver L., Bonnefont J.-P., Romano C., Flichera M.,
RA Lacombe D., Hanauer A., Philip N., Schwartz C.E., Fontes M.;
RT "Evaluation of a mutation screening strategy for sporadic cases of
RT ATR-X syndrome.";
RL J. Med. Genet. 36:183-186(1999).
[17]
RP VARIANTS ATR-X S-179; L-190; I-194; C-246; F-1552; S-1645 AND C-1847.
RX MEDLINE-20451413; PubMed-10995512;
RA Wada T., Kubota T., Fukushima Y., Saitoh S.;
RT "Molecular genetic study of Japanese patients with X-linked
RT alpha-thalassemia/mental retardation syndrome (ATR-X).";
RL Am. J. Med. Genet. 94:242-248(2000).
[18]
RP FUNCTION: COULD BE A GLOBAL TRANSCRIPTIONAL REGULATOR. MODIFIES
CC GENE EXPRESSION BY AFFECTING CHROMATIN. MAY BE INVOLVED IN
CC BRAIN DEVELOPMENT AND FACIAL MORPHOGENESIS.
CC -1- SUBUNIT: PROBABLY BINDS EZH2. BINDS ANNEKIN V IN A CALCIUM AND
CC PHOSPHATIDYLCHOLINE/PHOSPHATIDYLSELINE-DEPENDENT MANNER (BY
CC similarity).
CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH PERICENTROMERIC
CC HETEROCHROMATIN DURING INTERPHASE AND MITOSIS, PROBABLY BY
CC INTERACTING WITH HP1.
CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS: 1, 2, 3, 4 (SHOWN HERE) AND 5;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: UBQUITOUS.
CC DISEASE: DEFECTS IN ATRX ARE THE CAUSE OF X-LINKED ALPHA-
CC THALASSEMIA/MENTAL RETARDATION SYNDROME (ALSO KNOWN AS ATR-X
CC SYNDROME). ATR-X IS AN X-LINKED DISORDER COMPRISING SEVERE
CC PSYCHOMOTOR RETARDATION, FACIAL DYSMORPHISM, UROGENITAL
CC ABNORMALITIES, AND ALPHA-THALASSEMIA. AN ESSENTIAL PHENOTYPIC
CC TRAIT ARE HEMOGLOBIN H ERYTHROCYTE INCLUSIONS.

```

CC -1- DISEASE: DEFECTS IN ATRX ARE THE CAUSE OF CARPENTER-WIZIRI
CC SYNDROME (CWS), AN X-LINKED RECESSIVE CONDITION CHARACTERIZED BY
CC MODERATE MENTAL RETARDATION, SHORT STATURE, BRACHYACTYLIA WITH
CC EXCESSIVE SKIN CREASES, AND WIDENING OF THE KNUCKLES.
CC -1- DISEASE: DEFECTS IN ATRX ARE THE CAUSE OF JUBERG-MARSDI SYNDROME
CC (JM). JM IS A RARE X-LINKED RECESSIVE DISEASE CHARACTERIZED BY
CC SEVERE MENTAL RETARDATION, GROWTH FAILURE, SENSORINEURAL DEAFNESS,
CC MICROGENITALISM AND EARLY DEATH.
CC -1- DISEASE: DEFECTS IN ATRX ARE A CAUSE OF SMITH-FITENMAN-WYERS
CC SYNDROME (SFM). CLINICAL FEATURES INCLUDE SEVERE MENTAL
CC RETARDATION, MICROCEPHALY, GROWTH FAILURE, FACIAL ANOMALIES AND
CC BILATERAL CYCLOPHOSIDISM. DUE TO THE CLINICAL OVERLAP WITH ATR-X
CC SYNDROME, SOME PATIENTS ORIGINALLY DIAGNOSED AS HAVING SFM, MIGHT
CC BE AFFECTED BY A VARIANT OF ATR-X SYNDROME WHICH LACK HEMOGLOBIN H
CC INCLUSIONS.
CC -1- DISEASE: DEFECTS IN ATRX ARE THE CAUSE OF SUTHERLAND-HAAN X-LINKED
CC MENTAL RETARDATION SYNDROME (SHS). IT IS CHARACTERIZED BY SEVERE
CC MENTAL RETARDATION WITH SPASTIC PARAPLEGIA, MICROCEPHALY, SHORT
CC STATURE AND CRYPTORCHIDISM.
CC -1- SIMILARITY: BELONGS TO THE SMF2/RAD54 HELICASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PHD-TYPE ZINC FINGER.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U72937: AAB49970.2; -
DR EMBL: U72938: AAB49971.2; -
DR EMBL: U72935: AAB40698.1; -
DR EMBL: U72904: AAB40698.1; JOINED.
DR -----
Query Match 6.1%; Score 104; DB 1; Length 2492;
Best Local Similarity 18.7%; Pred. No. 11;
Matches 81; Conservative 54; Mismatches 142; Indels 156; Gaps 14;
CC 12 NPKSEYPMDC-----KKNINSHSGACMPRRKRLCYRDLTGGELIRKP 56
DB 671 NPYVSNSECENEYKEREKOLSVPRKKDRKSDSDALDNKPNKPKSKOSEYVONSND 730
CC 57 EDILTFINCAKETHFAHMKRYKKDNVAEN-----ELKSGKIPEGRKOMYYTFGDFR 110
DB 731 SDEMALILKGYSRMSSSSDPTDINEIHTNKTLYDLTKGAKDKKRRKSSSTGSGDF 790
CC 111 DIFPG-----TDISSRYIKDPSQTIK-SKLGDAATPK-----143
DB 791 DTRKGRSAKSIISKRRKQTSSESNVDSLEKEIKSKSIGAARTTKRIPNTRKDPSS 850
CC 144 -----GDTHI-----DNNKILQEWMTIHG-----PK 164
DB 851 EDEKSHKGMNQGKHNKLTQSGSSDAEKRQHEFTSSAGTYADKDTIMELRDRPK 910
CC 165 IWEGMICALJNGLSE-----SEKNIILODYSYNKLNAEKDDCLE 205
DB 911 --KQOASASTGVDLISGKEQSFSLAEYKVAETKEKSKHKTCKRKVQDGLSD--TAE 966
CC 206 KFAKSPQFLRWYEVMSDECFRERKLEDKVEDVCIAKADYEGCKNNKSNNSCVKCKEYE 265
DB 967 KFLAKDQ-----SDTSEDDKQSKKKGTEKKRPSDFK-----KVIKMEQOYE 1010
CC 266 NYITKK-----TOYEOGKFNTEKKRQKKEYSYSK-----299
DB 1011 SSSDTEKLPREBEICHFPKGIKQIKNGTIOSEKSKSKIRDTKSKKDELISYAKSKGK 1070
CC 300 ---DASEYLKDK 308
DB 1071 GDSCKSSDEKSKK 1083

```

```

YD33_SCHPO
ID YD33_SCHPO STANDARD: PRT: 278 AA.
AC Q10267;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C13G7.03 in chromosome I.
GN SPAC13G7.03.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkhardt G., Aert R., Robben J., Grynprenz B.,
RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Meller Auer S.,
RA Gabell C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Beer P., Zimmermann W., Medler H., Wambut R., Purrelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Talida V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Sipakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -1- SIMILARITY: SOME, TO HISTONES H1.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z69729; CAA93591.1; -
DR InterPro: IPR005120; Smg-4_UPEF3.
DR Pfam: PF03467; Smg4_UPEF3; 1.
KW Hypothetical protein.
SQ SEQUENCE 278 AA; 30935 MW; F3BCCCF71C29C85B CRC64;
Query Match 5.9%; Score 100; DB 1; Length 278;
Best Local Similarity 22.7%; Pred. No. 18;
Matches 64; Conservative 33; Mismatches 103; Indels 82; Gaps 13;
CC 56 PEDILTFINCAKETHFAHMKRYKKD--NVNAENELKS-----GKIPEFRKOMYY 104
DB 24 PEQYFLQINSFL--PVEVHFRFSKGRATGTSELSFAYLKFGOSATVAQVEFFRYQGH 81
CC 105 TFGDFRDIFFGTDISSCRYIKDTSQTIKSLGQATTEKEDTHIDNNKILQEWMTIHGPK 164
DB 82 TFDKNNTRYRAITYLAAYOKIIPSKVKA-----DSLEGSLEDDPKFOE-----K 127
CC 165 IWEGMICALJNGLSESEKKNIILODYSYNKLNN-----AEKDDCLEKFAKSPQFLRW 216

```

```

Db      128 VOR-----ESYSQTSASNDVDIEKLTPTPLLOYLAEKKNAVENKSKRP----- 177
Oy      217 YVENSDEGCRRRKLLEKVEDVCIAKADYBECCNNKNNSGVCKEKENITTKTOYE 276
Db      172 ----SKSVAKRKL-----RLAEKPASNNS--KAGSSQSKESSKAPAE 211
Oy      277 S-----OGKFNTERRRK-----PEXNSYSKKKASEYLKNC 308
|         | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      212 SAAAVIKEDKSDRKRKKRKKRKPVSNS-TASQASENASDK 252

RESULT 9
VDP_MOUSE STANDARD: PRT: 941 AA.
AC Q9Z120;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE General vesicular transport factor p15 (Transcytosis associated protein) (TAP) (Vesicle docking protein) (Fragment).
NE
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A. STRAIN=129; TISSUE=Thymus; Han S.J.;
"Identification of mouse TAP (transcytosis associated protein/p15)." ; Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
-i- FUNCTION: GENERAL VESICULAR TRANSPORT FACTOR REQUIRED FOR INTERCYTOSOLIC FUSION AND/OR SUBSEQUENT BINDING OF THE VESICLES TO THE TARGET MEMBRANE. MAY WELL ACT AS A VESICULAR ANCHOR BY INTERACTING WITH THE TARGET MEMBRANE AND HOLDING THE VESICULAR AND TARGET MEMBRANES IN PROXIMITY (BY SIMILARITY).
-i- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN WHICH RECYCLES BETWEEN THE CYTOSOL AND THE GOLGI APPARATUS DURING INTERPHASE (BY SIMILARITY).
-i- DOMAIN: COMPOSED OF A GLOBULAR HEAD, AN ELONGATED TAIL (COILED-COIL) AND A HIGHLY ACIDIC C-TERMINAL DOMAIN.
-i- PM: PHOSPHORYLATED IN A CELL CYCLE-SPECIFIC MANNER; PHOSPHORYLATED IN INTERPHASE BUT NOT IN MITOTIC CELLS. DEPHOSPHORYLATED PROTEIN ASSOCIATES WITH THE GOLGI MEMBRANE; PHOSPHORYLATION PROMOTES DISSOCIATION (BY SIMILARITY).
-i- SIMILARITY: BELONGS TO THE VDP/VSO1/YB1047C FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
Cc      EMBL; AF096868; AAC72967.1; .
DR MGD; MG1:1923095; VGP.
DR InterPro: IPR000225; Armaad110.
DR InterPro: IPR000219; RhogEF.
DR SMART; SM00325; RhogEF; 1.
KW PROSITE; PS50176; ARM_REPEAT; UNKNOWN_1.
KW Transport; Protein transport; Golgi stack; Membrane; Coiled coil; Phosphorylation.
FT FT NON_TER 1
FT FT DOMAIN <1 619 GLOBULAR HEAD.
FT FT DOMAIN 620 912 COILED COIL (POTENTIAL).
FT FT DOMAIN 917 941 ASP/GLU-RICH (ACIDIC).
FT MOD_RES 922 922 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 941 AA; 105152 MW; AB5CI0895CD7E508 CRC64;

Query Match 5.8%; Score 99.5; DB 1; Length 941;
Best Local Similarity 19.6%; Pred.No. 7.7;
```

| Matches | 52: | Conservative | 32: | Mismatches | 92: | Indels | 89: | Gaps | 6: |
|-----------|--|---|--|------------|-----|--------|-----|------|----|
| OY | 43 | CVRDITGGGELRKREDIL---- | TFINCAKETHFAMHKYKKDYNVAMNELSKSIPDGF | 98 | | | | | |
| Db | 460 | CNTLLSGGSKSIQTRVGLMLMLCTWLSNCPVIAVTHFLHN----- | SANVPELTQOIAE-- | 510 | | | | | |
| OY | 99 | RKQWYTFPGDPRDIFFGTDISSCRVIKQTSQTIKSLGDQATPTEKGQTHHDNNKKLQEMW | 158 | | | | | | |
| Db | 511 | ----- | NIGEEQVLQVG----- | 521 | | | | | |
| OY | 159 | TIHGPKIWEGMICALTNGLSESEKKNILQDYSYNNKLNNAEKDDCLEKFAASKPOPLRMVY | 218 | | | | | | |
| Db | 522 | ----- | LCALLGISIYFPNDNSLENTKREKLQILRKIGKEKFIKGLQFISKE | 570 | | | | | |
| OY | 219 | EWSD----- | FCREKKIKEDVEDYCIKAKYEGCKNNKSNNSCVKCKE | 263 | | | | | |
| Db | 571 | LYRSASQKQDPQFPSPPEYMIQDHEPTKLVKELEGVITKA--YKSSSEEDKKEEYKLTLEQ | 629 | | | | | | |
| OY | 264 | YENYITKKKTOYESECKFNTEKRO | 288 | | | | | | |
| Db | 630 | HDNVTYHKMIRREDQLQLEELKQO | 654 | | | | | | |
| RESULT 10 | | | | | | | | | |
| ID | YY02_METJA | STANDARD: | PRT: | 1181 | AA. | | | | |
| AC | 060301: | | | | | | | | |
| DT | 01-NOV-1997 (Rel. 35, Created) | | | | | | | | |
| DT | 01-NOV-1997 (Rel. 35, Last sequence update) | | | | | | | | |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update) | | | | | | | | |
| DE | Hypothetical protein MJEC502. | | | | | | | | |
| GN | MJEC502. | | | | | | | | |
| OS | Methanococcus jannaschii. | | | | | | | | |
| OC | Archaea; Euryarchaeota; Methanococci; Methanococcales; | | | | | | | | |
| OC | Methanocaldococcaceae; Methanocaldococcus. | | | | | | | | |
| OX | NCBI_TaxID=2190; | | | | | | | | |
| RN | 111 | | | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | | | |
| RC | STRAIN=JAL-1 / DSM 2661 / ATCC 43067; | | | | | | | | |
| RA | MEDLINE=96337999; PubMed=8688087; | | | | | | | | |
| RA | Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., | | | | | | | | |
| RA | Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D., | | | | | | | | |
| RA | Ketlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I., | | | | | | | | |
| RA | Overbeek R., Kirnke E.F., Weisskopf K.G., Merrick J.M., Glodek A., | | | | | | | | |
| RA | Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D., | | | | | | | | |
| RA | Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C., | | | | | | | | |
| RA | Cotton M.D., Roberts K.M., Hunt M.A., Kaine B.P., Borodovsky M., | | | | | | | | |
| RA | Klenk H.-P., Fraser C.M., Smith H.O., Woese G.C.R., Venter J.C.; | | | | | | | | |
| RT | "Complete genome sequence of the methanogenic archaeon, Methanococcus | | | | | | | | |
| RT | jannaschii."; | | | | | | | | |
| RL | Science 273:1058-1073(1996). | | | | | | | | |
| CC | -I- SIMILARITY: WEAK IN THE C-TERMINUS, TO ADENINE-SPECIFIC | | | | | | | | |
| CC | METHYLTRANSFERASES. | | | | | | | | |
| CC | ----- | | | | | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | | | | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | | | | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | | | | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | | | | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | | | | | | |
| CC | entities requires a license agreement (see http://www.isb-sib.ch/announce/ | | | | | | | | |
| CC | or send an email to license@isb-sib.ch). | | | | | | | | |
| CC | ----- | | | | | | | | |
| DR | EMBL: L77119; AAC37060.1; - | | | | | | | | |
| DR | TIGR: MJEC502; - | | | | | | | | |
| DR | InterPro: IPR002296; N12N6_mtfase. | | | | | | | | |
| DR | InterPro: IPR002052; N6_mfase. | | | | | | | | |
| DR | PRINTS: PR00507; N12N6MTFRASE. | | | | | | | | |
| DR | PROSITE: PS00092; N6_MTFASE; 1. | | | | | | | | |
| KM | Hypothetical protein: Transmembrane; Complete proteome. | | | | | | | | |
| FT | TRANSMEM 837 857 POTENTIAL. | | | | | | | | |
| SEQ | SEQUENCE 1181 AA; 138438 MW; 9F961D8F8CA4BDD CRC64; | | | | | | | | |

Query Match 5.8%; Score 99; DB 1; Length 1181;

Best Local Similarity 21.5%; Pred. No. 11;
Matches 63; Conservative 41; Mismatches 95; Indels 94; Gaps 13;

QY 41 KICVBDLTOGGEIRKPEDILKE-----INCAKETHFAMHY-----KQDNVNAEN 87
DB 8 KLC-----SDIEIPEEYNNKMKDFIKLVGNIIISIQIPELEKFEKLLKGRYNNRS 60
QY 88 ELKSGRIPEGF-RKQMYTFGDFRDIPEGTDISSCRYIKDTSQTIKSLGD----- 137
DB 61 DDFGQOLPEPTFRKQVIEPIEFLETFEFTSEK-----KSLPGDKKIDYKVS 109
QY 138 -----QATTEGDTHTIDNNK-----LOEMWTIHSPKIWEGMICALTNGLS----- 178
DB 110 VENKEILIEAPELGLSDLNKKDSCGHQVKEMLIKSYGVDTG--ATNGLEAVVLHYDDT 166
QY 179 -----ESEKKNILQDYSYNKLNNAEKDDCCLEK-----FASKPOFLRWYV----- 218
DB 167 IKEIRTLKLNKLSIFEYVLENNKDKD--LENEIKOVSEFYCYFSKYEIEYEVATK 223

219 -----EWSDECFERKKLEDDVEDV-CIRAKDEGCCKNNNSNVCKVE 263
224 NIKHKKEITNEFYKEFYKLVGFEDVDYKDKKDKSSSEKDKGTAKCLANCIE 276

RESULT 11
NUF1_YEAST
ID NUF1_YEAST STANDARD: PRT: 944 AA.
AC P32380:
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE NUF1 protein (spindle body protein SPC110).
GN NUF1 OR SPC110 OR YDR356W OR D9476.3.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RX MEDLINE=92176232; PubMed=1541631;
RA Mirzayan C., Copeland C.S., Snyder M.;
RT "The NUF1 gene encodes an essential coiled-coil related protein that
is a potential component of the yeast nucleoskeleton";
RL J. Cell Biol. 116:1319-1332(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94064779; PubMed=7503995;
RA Kilmarlin J.V., Dyos S.L., Keirshaw D., Finch J.T.;
RT "A spacer protein in the Saccharomyces cerevisiae spindle body
whose transcript is cell cycle-regulated.";
RL J. Cell Biol. 123:1175-1184(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Ravello A., Fulton L., Gattung S., Greco T., Kirsten J.,
RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA Marsden E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
RA Riffen L., Riles L., Taich A., Trevasakis E., Vignati D.,
RA Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEOSKELETON. POTENTIAL
IS ESSENTIAL FOR GROWTH.
CC -I- SUBCELLULAR LOCATION: NUCLEAR; TIGHTLY ASSOCIATED WITH THE
NUCLEUS. IT IS PRESENT IN A GRANULAR PATTERN THAT EXCLUDES THE
NUCLEOLUS.
CC -I- PTM: MAY BE REGULATED BY PHOSPHORYLATION EVENTS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: 211582; CAAT7668.1; -
DR EMBL: X73297; CAAS1733.1; -
DR EMBL: U28372; AAB64791.1; -
DR PIR: S26710; S26710.
DR PIR: S34288; S34288.
DR SCD: S0002764; NUF1.
KW Coiled coil; Nuclear protein; Phosphorylation.
FT DOMAIN 164 791
FT DOMAIN 54 59
FT DOMAIN 726 731
FT DOMAIN 742 747
FT DOMAIN 731 944
SQ SEQUENCE 944 AA; 111781 MW; 04FAA074B88A0BC8 CRC64;

Query Match 5.7%; Score 97.5; DB 1; Length 944;
Best Local Similarity 23.8%; Pred. No. 11;
Matches 81; Conservative 53; Mismatches 127; Indels 79; Gaps 20;

QY 21 DCKRNIDS-----HSGACMPRRQ---KLC-----VRDL-----TOGGEIRK 55
DB 208 DAKKKIVNSRKVDHSG-CIEERQMERKLAELERKILKTYVDYLENNSDVSLKRS 266
QY 56 PEDILTKI-----NCAKETHFAMHYKKKNVNAE-NELKSGRIPEFRQMYTF 106
DB 267 KEDELKLNMLNELKSAEERDQL--ERKKNELKRRTNELNELKIR----- 312
QY 107 GDFRIDPEGTDISSCRYIKDTSQTIKSL--GQATTEGDTHTIDNNK--KLOEMWTIH 161
DB 313 SDENDLQKQONKSKRLKDELNELFTKFSNGSSQSAKEMLKMLNKIALEELSTK 372
QY 162 GPKT--WEGMICALTNGISESEKKNILQDYSYNKLNNAEKDDCCLEKFAKPO---FLR 215
DB 373 NSOLAKEGKLASLMAQLTQLESKLNORD--SOLGSHREE--LKTNRNKLQKDIRAR 426
QY 216 WYVWSDECFER-RKKLEDDVEDVCIRAKDEGCCKNNKSN--NSCVACKREYEVYIGKK 272
DB 427 EETYSKDEIRIIDLOKKYKQLENDLFVIKTHSEKTTITDNELEKDKLIKLENDLVAQ 486
QY 273 TQYESQEGFENTERKQKPEYN---SYSK-KDASEYLKDK 308
DB 487 EKYSKM-----EKELKEREFNYKISSEKLEDEKTTLINEK 520

RESULT 12
SMC2_YEAST
ID SMC2_YEAST STANDARD: PRT: 1170 AA.
AC P38989;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-FEB-1995 (Rel. 31, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chromosome segregation protein SMC2 (DA-box protein SMC2).
GN SMC2 OR YFR031C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA Strunin A.V., Hogan E., Koshland D.;
RT "SMC2, a Saccharomyces cerevisiae gene essential for chromosome
segregation and condensation, defines a subgroup within the SMC
family";
RL Genes Dev. 9:587-599(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;


```

Db      283 GRSAVSPSEYDEEMIDILKDAGAEPMFTIPIVNGKMY-----321
QY      239 CIAKDEGCKNNKSNNSCVCKEYENYITGKKTYQESQEGKFNTEKROKKPEYNSYSK 298
Db      322 -----DVTGPKKGRFDYVKKVKNQ-----IRAKGFQVADFSGH-----EYDPEYM 362
QY      299 KD 300
Db      363 KD 364

```

RESULT 15

```

CYL2_HUMAN          STANDARD;          PRT;          348 AA.
ID CYL2_HUMAN
AC 014093:

```

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
RT 15-JUN-2002 (Rel. 41, Last annotation update)
CYL2_HUMAN II (Multiple-band polypeptide II).
CYL2 OR CYL2.

```

OC Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=95255491; PubMed=7737358.

RA Hess H., Heid H., Zimbelmann R., Franke W.W.;

RT "The protein complexity of the cytoskeleton of bovine and human sperm

heads: the identification and characterization of cyclin II.";

RL Exp. Cell Res. 218:174-182(1995).

CC -!- FUNCTION: POSSIBLE ARCHITECTURAL ROLE DURING SPERMATOGENESIS. MAY

BE INVOLVED IN SPERMATID DIFFERENTIATION.

CC -!- SUBCELLULAR LOCATION: CALYX; SPERM HEAD CYTOSKELETAL STRUCTURE.

CC -!- TISSUE SPECIFICITY: TESTIS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL: Z46788; CA86752.1; -

Genew; HGNC:2583; CYL2.

MIM; 604035; -

Cytoskeleton; Structural protein; Repeat; Sperm; Spermatogenesis.

FT DOMAIN 25 347 31 X 3 AA REPEATS OF K-K-X.

FT REPEAT 157 240 3 X APPROXIMATE TANDEM REPEATS.

FT REPEAT 157 184 1.

FT REPEAT 185 212 2.

FT REPEAT 213 240 3.

SEQUENCE 348 AA; 39079 MW; D86766599C1809E7 CRC64;

Query Match 5.7%; Score 96.5; DB 1; Length 348;

Best Local Similarity 23.8%; Pred. No. 4.2;

Matches 72; Conservative 33; Mismatches 104; Indels 93; Gaps 13;

```

QY      38 RROKL-----CVRDLTGGEL-RRPEDILTFKFINCAKETHFAWHKRYKKDNVNAENELKSG 92
Db      95 RROPLKPTFIVEVDSKAELGKGEDKT-----OKDTTDSSEELKOG 137
QY      93 KIPGFRKQMYVFGDFRIDFPGTDISSCRYIK-DTSQTIK-----SKLGDQAT---TE 142
Db      138 K-----KDSKKGKIDIEKGEKELDAKKDSKKGKDAEKDASATESEDE 181
QY      143 KGDTHEDNKKLOEMWFIHGPKIWEGLCALITNGL-----SESKNIILODYSYNKLNNA 197
Db      182 KGAKKDKKKDKKD-----SNKKGDSATESEGEGGTGTERKSKKGGKDKSK 225

```

```

QY      198 EKDDCCLEKFPASRPOFLRMVYEWSDFCRRRKKLEDKVEDVCIKAKDYECKNNKSNNSC 257
Db      226 KGDASATIELQAVKAD-----EKKDEGCKKDKANKGDSKDAKKDAKIRKGGKDKKKPSS 279
QY      258 V-----KCKEYE-----NYTGKRTYESQEGKFNTEKROKKPEYNSYSKQASEY 304
Db      280 TDSKSDVKKESKDKATKDAKKVAKKDTESADSKKDAKKNAKND-----AKKDAKKN 334
QY      305 LK 306
Db      335 AK 336

```

Search completed: June 20, 2003, 15:03:34
Job time : 16.5539 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 14:54:09 ; Search time 48.1396 Seconds
(Without alignments)
1318.301 Million cell updates/sec

Title: US-10-087-013-11

Perfect score: 1706

Sequence: 1 GNDGSENEISGNCNPKESYPDW.....KKREYNSYSKKDASEYLKDK 308

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

al number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.protist:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.rvirts:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 1706 | 100.0 | 2710 | 5 | Q9XZB8 |
| 2 | 894 | 52.4 | 3006 | 5 | Q26032 |
| 3 | 669.5 | 39.2 | 3078 | 5 | Q26031 |
| 4 | 617.5 | 36.2 | 3287 | 5 | Q8T326 |
| 5 | 600 | 35.2 | 3026 | 5 | Q26030 |
| 6 | 572 | 33.5 | 3542 | 5 | Q9U5M2 |
| 7 | 571.5 | 33.5 | 2658 | 5 | Q8T5G0 |
| 8 | 569.5 | 33.4 | 1615 | 5 | Q8T325 |
| 9 | 526 | 30.8 | 2706 | 5 | Q15870 |
| 10 | 521 | 30.5 | 921 | 5 | Q25989 |
| 11 | 515.5 | 30.2 | 2527 | 5 | Q95W83 |
| 12 | 444 | 26.0 | 1685 | 5 | Q9U4A2 |
| 13 | 437 | 25.6 | 2209 | 5 | Q9U0G6 |
| 14 | 433 | 25.4 | 1711 | 5 | Q96108 |
| 15 | 407.5 | 23.9 | 2169 | 5 | Q97312 |
| 16 | 395 | 23.2 | 2042 | 5 | Q25766 |

| | | | | | |
|----|-------|------|------|---|--------|
| 17 | 386 | 22.6 | 2228 | 5 | Q60991 |
| 18 | 385.5 | 22.6 | 2277 | 5 | Q9U0G5 |
| 19 | 383 | 22.5 | 2212 | 5 | Q94657 |
| 20 | 375 | 22.0 | 184 | 5 | Q9BUE6 |
| 21 | 371.5 | 21.8 | 2209 | 5 | Q97324 |
| 22 | 368.5 | 21.6 | 195 | 5 | Q00831 |
| 23 | 367.5 | 21.5 | 2182 | 5 | Q26034 |
| 24 | 362 | 21.2 | 176 | 5 | Q9BJR0 |
| 25 | 361.5 | 21.2 | 2647 | 5 | Q90580 |
| 26 | 354 | 20.8 | 2135 | 5 | Q61077 |
| 27 | 354 | 20.8 | 2163 | 5 | Q9NEB6 |
| 28 | 352 | 20.6 | 2664 | 5 | Q26033 |
| 29 | 342.5 | 20.1 | 2197 | 5 | Q96296 |
| 30 | 333 | 19.5 | 182 | 5 | Q00832 |
| 31 | 327.5 | 19.2 | 616 | 5 | Q8T9N3 |
| 32 | 327.5 | 19.2 | 616 | 5 | Q8T9L8 |
| 33 | 327.5 | 19.2 | 1210 | 5 | Q96774 |
| 34 | 327.5 | 19.2 | 1210 | 5 | Q8WS31 |
| 35 | 325.5 | 19.1 | 177 | 5 | Q9BUE8 |
| 36 | 321 | 18.8 | 162 | 5 | Q9BUE7 |
| 37 | 311 | 18.2 | 2924 | 5 | Q25733 |
| 38 | 298.5 | 17.5 | 1729 | 5 | Q25734 |
| 39 | 286 | 16.8 | 173 | 5 | Q25988 |
| 40 | 268.5 | 15.7 | 699 | 5 | Q25987 |
| 41 | 265.5 | 15.6 | 2647 | 5 | Q9U4X0 |
| 42 | 265 | 15.5 | 420 | 5 | Q9BHB9 |
| 43 | 265 | 15.5 | 420 | 5 | Q9BHB2 |
| 44 | 263 | 15.4 | 169 | 5 | Q25990 |
| 45 | 263 | 15.4 | 420 | 5 | Q9BMY4 |

ALIGNMENTS

RESULT 1
Q9XZB8 PRELIMINARY: PRT: 2710 AA.

AC Q9XZB8: 01-NOV-1999 (TREMURel. 12, Created)
DT 01-NOV-1999 (TREMURel. 12, Last sequence update)
DT 01-DEC-2001 (TREMURel. 19, Last annotation update)
DE Variant-specific surface protein.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
CX NCBI_TaxID=5633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS2;
RX MEDLINE=99238507; PubMed=10220443;
RA Rogers J.C., Cowman A.F., Davern K.M., Beeson J.G., Thompson J.K.,
RT "The adhesion of Plasmodium falciparum-infected erythrocytes to
RT chondroitin sulfate A is mediated by P. falciparum erythrocyte
RL Proc. Natl. Acad. Sci. U.S.A. 96:5198-5202(1999).
DR EMBL: AF134154; AAD29126.1;
DR InterPro: IPR004258; PfEMP.
DR Pfam: PF03011; PfEMP.
SQ SEQUENCE 2710 AA; 309247 MW; 0AC64F0A3D5BF512 CRC64;

Query Match 100.0%; Score 1706; DB 5; Length 2710;
Best Local Similarity 100.0%; Pred. No. 1,je-115;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNDGSENEISGNCNPKESYPDWCKRNINDNSHGACMPRRKRLCYRDLTGGELRRKPEDIL 60
Db 908 GNDGSENEISGNCNPKESYPDWCKRNINDNSHGACMPRRKRLCYRDLTGGELRRKPEDIL 967
QY 61 TKFLNCAKETHFMHKKRNVAENELKSGKIPGFRKMYTTTFGDFRIFGCTDISS 120
Db 968 TKFLNCAKETHFMHKKRNVAENELKSGKIPGFRKMYTTTFGDFRIFGCTDISS 1027

```

QY 121 CRYIKDTSOTIKSLGDQATEKGDTHIDNKKLOEWMTIHGKIEWGMICALTNLSSES 180
DB 1028 CRYIKDTSOTIKSLGDQATEKGDTHIDNKKLOEWMTIHGKIEWGMICALTNLSSES 1087
QY 181 EKKNILODYSYNKLNNAEKDDCCLEFASKPOFLRMVEMSDSEFCERKLEKVDVCI 240
DB 1088 EKKNILODYSYNKLNNAEKDDCCLEFASKPOFLRMVEMSDSEFCERKLEKVDVCI 1147
QY 241 KADYEGCKNNKSNNSCVAVCKEYENYITGKTQYESQEGKFTERKOKKPEYNSYKRD 300
DB 1148 KADYEGCKNNKSNNSCVAVCKEYENYITGKTQYESQEGKFTERKOKKPEYNSYKRD 1207
QY 301 ASFYLKDK 308
DB 1208 ASFYLKDK 1215

```

RESULT 2

```

Q26032 PRELIMINARY; PRT; 3006 AA.
ID 026032; 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
DE Variant-specific surface protein.
RT cytoadherence and antigenic variation of Plasmodium falciparum-
RN Cell 82:89-100(1995).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=FCR3;
RX MEDLINE=95330813; PubMed=7606788;
RA Su X.Z., Heatwole V.M., Wertheimer S.P., Guinet F., Herfeltdt J.A.,
RA Peterson D.S., Ravetch J.A., Wellens T.E.;
RA "The large diverse gene family var encodes proteins involved in
RT cytoadherence and antigenic variation of Plasmodium falciparum-
RT infected erythrocytes."
RL Cell 82:89-100(1995).
DR EMBL; L40609; AAA75397.1;
DR InterPro; IPR004258; PFEMP.
DR Pfam; PF03011; PFEMP. 2.
SQ SEQUENCE 3006 AA; 343769 MW; 8FDSFF475F943C74 CRC64;

```

Query Match 52.4%; Score 894; DB 5; Length 3006;
 Best Local Similarity 53.0%; Pred. No. 1.7e-56;
 Matches 172; Conservative 38; Mismatches 91; Indels 12; Gaps 4;

```

5 SNEISGCP-----KESYPMDCCKNIDNSHGACMPPRROKLCVRLTGGERTKRPDI 59
1236 TDDIDGCKQKAKGDKIPGMDCSQIHTHNGACMPPRROKLCVSLTKTDRIKATIEYI 1295
QY 60 LTRKFLNCAKETHPAMHKYKKDNVAENELKSGKIPGFRKQMYTFGDFRDIFFGTDIS 119
DB 1296 RTEFIKSAIETHPAMDYKEDNGEAEALKNGINPEGRKQMYTFGDFRDIFFGTDIS 1355
QY 120 SCRYIKDTSOTIKSLGDQATEKGDTHIDNKKLOEWMTIHGKIEWGMICALTNNG 176
DB 1356 THAVISGVSPKVTITLEKENAKYAAKQNS--NNEILDMDWDHGXKIMEGMICALTHK 1412
QY 177 LSESE-KKNILQDYSYNKLNNAEKDDCCLEFASKPOFLRMVEMSDSEFCERKLEKVDV 235
DB 1413 ISDEKKEKEIKNNKSYKLNKSPGSKNVEFAKPPQFLRMVEMSDSEFCERKLEKVDV 1472
QY 236 EDVCITAKDYEGCKNNKSNNSCVAVCKEYENYITGKTQYESQEGKFTERKOKKPEYNS 295
DB 1473 KVSCKDAKDYDGCNKTNSNASCYACVYEDYITKKVEYETKQKGTDAEKITTEGEGEG 1532
QY 296 YSKKDASEYLLKDK 308
DB 1533 FSTKDASEYLLKDK 1545

```

RESULT 3

```

Q26031 PRELIMINARY; PRT; 3078 AA.
ID 026031; 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
DE Variant-specific surface protein.
RN Cell 82:89-100(1995).
OS Plasmodium falciparum (isolate Dd2).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57267;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=DD2;
RX MEDLINE=95330813; PubMed=7606788;
RA Su X.Z., Heatwole V.M., Wertheimer S.P., Guinet F., Herfeltdt J.A.,
RA Peterson D.S., Ravetch J.A., Wellens T.E.;
RA "The large diverse gene family var encodes proteins involved in
RT cytoadherence and antigenic variation of Plasmodium falciparum-
RT infected erythrocytes."
RL Cell 82:89-100(1995).
DR EMBL; L40608; AAA75396.1;
DR InterPro; IPR004258; PFEMP.
DR Pfam; PF03011; PFEMP. 2.
SQ SEQUENCE 3078 AA; 349297 MW; C8037C2B3C3CD7C3 CRC64;

```

Query Match 39.2%; Score 669.5; DB 5; Length 3078;
 Best Local Similarity 40.9%; Pred. No. 3.8e-40;
 Matches 133; Conservative 54; Mismatches 109; Indels 29; Gaps 9;

```

QY 1 GNDGSNEISGCPKPSYPMDCCKNIDNSHGACMPPRROKLCVRLTGGERTKRPDI 57
DB 1356 GKNSTTGTGECNPKPSYPMDCCKNIDNSHGACMPPRROKLCVRLTGGERTKRPDI 1415
QY 58 DILTRKFLNCAKETHPAMHKYKKDNVAENELKSGKIPGFRKQMYTFGDFRDIFFGTDIS 117
DB 1416 NIKDAFIKTAAEFTLSWQYKSKNDSSEAKITLDGLIPSGFLRSMVTFGDFRDIFFGTDIS 1475
QY 118 ISSCRYIKDTSOTIKSLG-----DQATEKGDTHIDNKKLOEWMTIHGKIEWGMICAL 172
DB 1476 ISKRO--NDVAKA-KDKIGKFSKDGSKSPSLR-----QEMWKTNGPEIKWGMICAL 1525
QY 173 LTNGLSESE-KKNILQDYSYNKLNNAEKDDCCLEFASKPOFLRMVEMSDSEFCERKLEKVDV 231
DB 1526 LTRKVTLDNKKRKIKNDYSYKLVQSONGNLFPAKPPQFLRMVEMSDSEFCERKLEKVDV 1585
QY 232 EDKVEDVCITAKDYEGCKNNKSNNSCVAVCKEYENYITGKTQYESQEGKFTERKOKK 289
DB 1586 ENIKIDACNEINSTQOC--NDAKHRCQACRAVOGEYENKKEKFSGOTNNFVLKANVOPQ 1643
QY 290 KPEYNSYKRD-----ASEYLLKDK 308
DB 1644 DPEYKGYEYKDGVOPIQDNEYLLQK 1668

```

RESULT 4

```

O8T326 PRELIMINARY; PRT; 3287 AA.
ID 08T326; 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, last sequence update)
DE Variant-specific surface protein.
RN TMT80VAR2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=TM180;
RX MEDLINE=21927235; PubMed=11930336;
RA Rowe J.A., Kyes S.A., Rogerson S.J., Babiker H.A., Raza A.;

```

RT "Identification of a conserved Plasmodium falciparum var gene implicated in malaria in pregnancy."
 RL J. Infect. Dis. 185:1207-1211(2002).
 DR EMBL: AJ420411; CAD20867.1; -
 FT NON_TER 3287
 SQ SEQUENCE 3287 AA; 383550 MW; 588F866FC244536 CRC64;
 Query Match 36.2%; Score 617.5; DB 5; Length 3287;
 Best Local Similarity 41.7%; Pred. No. 2,5e-36;
 Matches 130; Conservative 56; Mismatches 101; Indels 25; Gaps 11;
 QY 8 ISGCPKRS--YPMDCCKNIDNS-HSGACMPRRQKLCVRL--TGGGEIRKE-DILT 61
 Db IEGCNPRTNGKYPKMEC--NVCKSKENACMPRRKIKCINNIQYLNKTEKPEKIDKN 2138
 QY 62 KFINCAKEHPHAMKYYKKNVNAENELSKIPGCFKQMYTTCGDRDIFGDISC 121
 Db AFICAALEIFYFWLYKLDNPTAENELKNGRIPEEFKIMYYTGYDKDMFEFGDITNE 2198
 122 RYIKDSOTIKSKLDQATTEKGDPIHIDNKKIQEMWTIHGPKIMEGMLCALTNGLS-ES 180
 2199 SKIITINNNTIILKNENKKKQDKRTDELLKNFWEENKKLIMEGMITGLTHLTDET 2258
 QY 181 EKKNILODYSTINKLNAEKDDCCLEKFAKPOFLRWYEWSDFCRERKKLEKVEDYCI 240
 Db EKKKTTDSEYTDIKLSPS--LEEFARPOFLRWFTEMAEFCKRKK--OLENLKE 2312
 QY 241 KAKDEGCKNNKNSCYKVCKEYENYITGKKTOYESOGKFNTEKROKKPEYNSY---- 296
 Db 2313 KCPDVT-CSVDSTKKECDSCAKYOTFIKQMTQYERKGRKNDKDTDK--YKDIPLIS 2369
 QY 297 ---SKKASEYL 305
 Db 2370 YIDNATDAHEYL 2381
 RESULT 5
 Q26030 PRELIMINARY; PRT; 3026 AA.
 AC Q26030;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Variant surface protein (Fragment).
 GN VAR.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NCBI_TaxId=5833;
 (11)
 SEQUENCE OF 2044-2922 FROM N.A.
 RC STRAIN=IT 4/25/5;
 RA MEDLINE=95330813; PubMed=7606788;
 RA Su X.Z., Heatwole V.M., Wertheimer S.P., Guinet F., Herfelldt J.A., Peterson D.S., Ravetch J.A., Wellens T.E.;
 RA "The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes."
 RT Cell 82:89-100(1995).
 RL (12)
 SEQUENCE OF 2044-2922 FROM N.A.
 RC STRAIN=IT 4/25/5;
 RA Smith J.D., Chlitis C.E., Craig A.G., Roberts D.J., Hudson-Taylor D.E., Peterson D.S., Plinches R., Newbold C.I., Miller L.H.;
 RA "Switches in the expression of Plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes."
 RT Cell 0:0-0(1995).
 RL (13)
 SEQUENCE FROM N.A.
 RC STRAIN=IT 4/25/5;
 RA MEDLINE=99094502;
 DR Smith J.D., Kyes S., Craig A.G., Fagan T., Hudson-Taylor D.,

RA Miller L.H., Baruch D.I., Newbold C.I.;
 RT "Analysis of adhesive domains from the AVAR Plasmodium falciparum erythrocyte membrane protein-1 identifies a CD36 binding domain."
 RL Mol. Biochem. Parasitol. 97:133-148(1998).
 RN (14)
 RP SEQUENCE FROM N.A.
 RC STRAIN=IT 4/25/5;
 RA Smith J.;
 RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
 RN (15)
 RP SEQUENCE FROM N.A.
 RC STRAIN=IT 4/25/5;
 RA Kyes S., Smith J.;
 RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: I42244; AA03351.1;
 DR InterPro: IPR004258; PfEMP.
 DR Pfam: PF03011; PfEMP. 2.
 FT NON_TER 3026
 SQ SEQUENCE 3026 AA; 348124 MW; EC0A971A82193788 CRC64;
 Query Match 35.2%; Score 600; DB 5; Length 3026;
 Best Local Similarity 35.9%; Pred. No. 4.3e-35;
 Matches 122; Conservative 58; Mismatches 112; Indels 48; Gaps 9;
 QY 1 GNDGSEISGCGN-----PSEYPMDCCKNIDNSHSGACMPRRQKLCVRLTGGGEIRK 55
 Db 1996 GGGDTKRIIECQTKRYPTKNDYFGNNCTDKVINREGSCMPRRQKLCIHNLHLSKAT 2055
 QY 56 PEDILTKFINCAKETHEFMHKKYKKNV-----NAENELSKIPGCFK 100
 Db 2056 ETELKRAFLTECAAIETFWLMDKYEKDEKTEGGGIGSDPDDPOKKLEGGTIPEDFKR 2115
 QY 101 QMYTFGDFRDIFFGDISSCXYINDTSOTIKSKLDQAT--EKGDPIHIDNKKIQEMW 158
 Db 2116 QMFYTYGDRDFLFGTIDSK-----GCKESALGKKTIDSLFKNGDQSPSGKPTPEW 2168
 QY 159 TTHGPKIMEGMLCALT---NGLSESEKKNILQDYSYNKLNAEKDDCCLEKFAKPOFL 214
 Db 2169 NNYGPDIMKGMVCGLSHHIKNKNGKIDQLRKNLTDNNKYTISSK-----LEDFASPOFL 2222
 QY 215 RWYVMSDFECGERKKLEKVEDYCIKADYE-GCKNNKSNNSCYKVCKEYENYITGKKY 273
 Db 2223 RWFIEWGDOFCREYRVKINQLTGC---NEYEGSGSENGKCKACKNACPAKRWLKDWD 2279
 QY 274 QYESOGKFNTEKROK-----PEYNSYKKNASEYLKMK 308
 Db 2280 QYEQQTAKFDKDKKDKKEDGTSAEVDAVAVSSVHEYLDE 2319
 RESULT 6
 Q905M2 PRELIMINARY; PRT; 3542 AA.
 AC Q905M2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE FCGR3 CSA ligand (Fragment).
 GN VAR.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NCBI_TaxId=5833;
 (11)
 SEQUENCE FROM N.A.
 RC STRAIN=FCGR3;
 RA MEDLINE=20006305; PubMed=10535993;
 RA Butfel P., Gamain B., Scheidig C., Baruch B., Oishi S., Fujii N., Fusai T., Parzy D., Miller L.H., Gysin J., Scherf A.;
 RA "Plasmodium falciparum domain mediating adhesion to chondroitin sulfate A: A receptor for human placental infection."
 RT Proc. Natl. Acad. Sci. U.S.A. 96:12743-12748(1999).
 RL EMBL: AJ13811; CAB59840.1; -
 DR InterPro: IPR001219; Neurotoxin.
 DR InterPro: IPR004258; PfEMP.

DR Pfam; PF03011; PFEMP; 1.
 DR PRINTS; PRO0284; TOXIN.
 FT NON_TER 3542 3542
 SO SEQUENCE 3542 AA; 413089 MW; 970DB5EB8DA2EC2 CRC64;

Query Match 33.5%; Score 572; DB 5; Length 3542;
 Best Local Similarity 39.9%; Pred. No. 5.6e-33;
 Matches 131; Conservative 54; Mismatches 105; Indels 38; Gaps 14;

QY 3 DGSNLSGCPNPK--SYPDWDC--KKNINSHSGACMPRRKRLCYRD---LTQGEIRK 55
 DB 2092 DSRGEGEGCPNPKYGOYPRKGCIVGSKEN--ENGICMPRRKRLCTINNIOYLVEYENKR 2150

QY 56 PEDILKFTINCAKAEHFAMHKKYKKNVNAENELKSGKIPRGRKOMYTFGDFRDIFFG 115
 DB 2151 DNDIKAFIKCAIEIETQFLMLKTYIENPAENLQGTIDPEKRLIMYYGYGKMFEG 2210

QY 116 TDISSCRVYKDTQSQTISKIGDQATTEKGTIDNNKLOE---KWTTHGPKIMEGMLC 171
 DB 2211 TDISNDRKIIITVNSVTTLINENKRRQ-----DKKDELRKIFWEKKKFIWEGMITY 2264

QY 172 ALTNGLS--ESEKKNIIQDYSYNKLNNAEKDDCLEKFAKQPLRRYVENSDEFCRRK 230
 DB 2265 GLTYHLTDENKERRKIDNTQYNDMTKLTPS---LEEFVRRPQLRMTFTEAEFCCKRRE 2321

QY 231 LEDKVEDVICAKDYEGC--KNNKSNNSCVKVEKENYITGKRTOYESQEGFENREKQ 288
 DB 2322 QLLKLKAGC---KEYE-CNNSNDGKTOEGCAECVTONFLKKMKTEYERREKFKDKOC 2377

QY 289 KRPEVNSY--SKKD-----ASFYLLKDK 308
 DB 2378 KK--YKDYPTSTERDIEKATCAHEYLNMK 2403

RESULT 7
 ID 08T5G0 PRELIMINARY; PRT; 2658 AA.
 AC 08T5G0;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE Erythrocyte membrane protein-1 (Fragment).
 GN VAR.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 PP SEQUENCE FROM N.A.
 STRAIN=JDP8;
 RT "Identification of a domain responsible for binding to intercellular
 adhesion molecule-1 from a Plasmodium falciparum field isolate."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ028643; AAK49742.1; -;
 FT NON_TER 2658 2658
 SO SEQUENCE 2658 AA; 301326 MW; AEE33BC9167E7B7 CRC64;

Query Match 33.5%; Score 571.5; DB 5; Length 2658;
 Best Local Similarity 38.0%; Pred. No. 4.4e-33;
 Matches 127; Conservative 54; Mismatches 102; Indels 51; Gaps 12;

QY 2 NDGSNEISGCPNPK-----SYPDWDC--KKNINSHSGACMPRRKRLCYRDLTQGEIRK 55
 DB 1317 DDBTKGVGCGNAIKININSEYIPMTVCVNSFENNEGCPMPRRKRLCTLYLKELEND 1376

QY 56 PEDILKFTINCAKAEHFAMHKKYKKNVNAENELKSGKIPRGRKOMYTFGDFRDIFFG 115
 DB 1377 EOKFKDAFIKTAFAETFLSQYKSKNSMDIKLQSGEIPRGRMYFYGYDRDLCN 1436

QY 116 TDISSCRVYKDTQSQTISKIGDQATTEKGTIDNNKLOEWTTHGPKIMEGMLC 169
 DB 1437 TDIS-----KKEGD--VSDAKGKIDAFKNTYDITNR--TKWMDTNGPEIWEGM 1480

QY 170 LCATLNGLESE--KKNIIQDYSYNKLNNAEKDDCLEKFAKQPLRRYVENSDEFCRRK 228
 DB 1481 LCATLHGVINTDNRKRIKIDYSYKL--QSKNVTTPLEKFAERQFLRMILEMGDDDCRRQ 1539

QY 229 KRIEDVEDVCYKADYEGCKNNKSNNS-----CYKVEKENYITGKRTOYESQEG 279
 DB 1540 KKKYNELKERCNK-----CNNGNVTSDCKTRVCECKRQKGEYKGTTEWQENMKOR 1593

QY 280 GRFNEKRO--KRPEVNSYKKAASE-----LYKD 307
 DB 1594 NKETLYQVKTSTSRSTISSDPTETKLKLYNE 1627

RESULT 8
 ID 08T325 PRELIMINARY; PRT; 1615 AA.
 AC 08T325;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE pEEM1 (Fragment).
 GN TM284VAR3.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 PP SEQUENCE FROM N.A.
 STRAIN=TM284;
 RT MEDLINE=21927235; PubMed=11930336;
 RA Howe J.A., Kyes S.A., Rogerson S.J., Babiker H.A., Raza A.;
 RT "Identification of a conserved Plasmodium falciparum var gene
 implicated in malaria in pregnancy."
 RL J. Infect. Dis. 185:1207-1211(2002).
 DR EMBL: AJ420412; CAD20868.1; -;
 FT NON_TER 1615 1615
 SO SEQUENCE 1615 AA; 186639 MW; FCA896C00B86DEA6 CRC64;

Query Match 33.4%; Score 569.5; DB 5; Length 1615;
 Best Local Similarity 37.3%; Pred. No. 3.6e-33;
 Matches 125; Conservative 57; Mismatches 104; Indels 49; Gaps 14;

QY 1 GNDGSNEISGCPNPK--SYPDWDC--KKNINSHSGACMPRRKRLCYRDLTQGEIRK 58
 DB 1253 GNGGTTKVGECNVKTEGEYRPEWCEKRIQNIHGAACMPRRKRLCYHFLKE--LKVETDER 1311

QY 59 ILTKFTINCAKAEHFAMHKKYKKNVNAE--NELKSGKIPRGRKOMYTFGDFRDIFFG 115
 DB 1312 LREAFIQCSAETFLMKRYKEDNNGEDLQNLQESGILIPDCKROMETTFGDRDLG 1371

QY 116 TDISSCRVYKDTQSQTISK--KLQDQATTEKGTIDNNKLOEWTTHGPKIMEGMLC 172
 DB 1372 KDIGS--DVGNVNAKINSALKTIGQ-----SDEKRRNMWNGIKEDYKMGVCG 1418

QY 173 LINGLESER--KNIID-----YSYNKLN--NAEKDDCLEKFAKQPLRRYVENS 222
 DB 1419 LSHDVSGNHKEIYRKRKIMEDPONNKYQVNSVKTDEPSTGKISEFAKVQPLRMLTEWMD 1478

QY 223 EFCREKRLKEDVEDVCIAKADYEGCKNNKSNNSCVKVEKENYITGKRTOYESQEG 282
 DB 1479 DYCHTQKTLKEVESTC--KSNQDLK-----DTECKKCEDEYKMKKKKEWIIOD--KY 1531

QY 283 NTEKROKPEYN-----SYSKKDASEYLLKDK 308
 DB 1532 YKDERDK--RFNRQHGVAWKDYTGINAIDYLLRK 1565

RESULT 9
 ID 015870 PRELIMINARY; PRT; 2706 AA.
 AC 015870;
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)

```

DE PFEMP1 (Fragment).
GN R29R-VARI.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=5833;
RN
  [1]
  SEQUENCE FROM N.A.
  RP STRAIN=IT 4/25/5;
  RX MEDLINE=97373957; PubMed=92230440;
  RA Rowe J.A., Moulds J.M., Newbold C.T., Miller L.H.;
  RT "Plasmodium falciparum rosetting is mediated by PFEMP1 and requires
  complement receptor 1."
  RL Nature 388:292-295(1997).
  DR EMBL: Y13402; CAAT3831.1; -.
  DR EMBL: Y13403; CAAT3831.1; JOINED.
  DR InterPro: IPR002086; Aldehyde_dehydr.
  DR InterPro: IPR004258; PFEMP.
  DR Pfam: PF03011; PFEMP; 2.
  DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
  DR NON_TER 2706 2706
  SEQUENCE 2706 AA; 308162 MW; 1C33D55AD5317D68 CRC64;

Query Match 30.8%; Score 526; DB 5; Length 2706;
Best Local Similarity 36.6%; Pred. No. 9.3e-30;
Matches 117; Conservative 49; Mismatches 122; Indels 32; Gaps 12;

QY 5 SNEISGCPNK-ESYPMDCCKN-IDNSHSGACMPRRORCLVADLTGSGEIRKPEDILTK 62
DB 782 NSGIDCNMKNRRKNEMOCDKNTFVGNESVCMPPRKISICINLLEOTKKNYQUREA 841
QY 63 FINCARETHFAWHKRYKDVNAENELSGKIEGFRKQMYTFGDFRIFETGDISCR 122
DB 842 FICCAKETLMDKYNKKNENAEELKKGKIPEDFRIRMFYFGDFRDLCEMDNG---898
QY 123 YIDTSQTIKSKIGD--QATTEKGDTHIDNKKLQEMWTIHGKPIEGMICALTNGLS 180
DB 899 --KDVK--VKKNINKVFNNSSKRGFKIDP---EWMWNGPOINGMICALIHADTK 951
QY 181 EKKNIIDQSYNKLNNAEKDD---CCLKEFASKPOFLRYVMSDFEGRERKLEKVE 236
DB 952 SIN-KDNKYEEVYTLAKDGSNGMTLSEFAKPKFLRFVWYDYCKERKYLTEVA 1010
QY 237 DVCIRAKDEGCKNNKSNNSCVKVEYENYITGKTOYSGEGRNTEKRRKP-----291
DB 1011 STC---KSIDG--QLMCDRCGCKNCKDEYKYMRRKKKEWMLQD-KYKKDKREKKGIDKP 1065
QY 292 ---EYNSYKSKDASEYLKDK 308
DB 1066 IGIIIVKDYVLNAKEYLKK 1085

RESULT 10
Q25989 PRELIMINARY; PRT; 921 AA.
AC Q25989;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE (clone PS31H) ORF (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=5833;
RN
  [1]
  SEQUENCE FROM N.A.
  RP STRAIN=DD2/NM;
  RX MEDLINE=95350219; PubMed=7624377;
  RA Peterson D.S., Miller L.H., Wellem T.E.;
  RT "Isolation of multiple sequences from the Plasmodium falciparum genome
  that encode conserved domains homologous to those in erythrocyte-
  binding proteins."
  RL Proc. Natl. Acad. Sci. U.S.A. 92:7100-7104(1995).
  DR EMBL: L38454; AAC37240.1; -.
  DR InterPro: IPR004258; PFEMP.

```

```

DR Pfam: PF03011; PFEMP; 1.
FT NON_TER 1
SQ SEQUENCE 921 AA; 104711 MW; D98BC07763628746 CRC64;

Query Match 30.5%; Score 521; DB 5; Length 921;
Best Local Similarity 37.9%; Pred. No. 6.6e-30;
Matches 127; Conservative 46; Mismatches 106; Indels 56; Gaps 13;

QY 3 DGSNEISGCPNKESYPMDC-KKNIDNSHSGACMPRRORCLVADLTGSGEIRKPEDILTK 61
DB 407 NGRTTVEGCEKRYSEMTDESKIKMGQHGACIPPRORCLHYL-----EKMT 457
QY 62 K-----FINCAKETHEFAWHKRYKDVNAEN---ELKSGKIEGFRKQMYTFGDFR 110
DB 458 NTMELKAFATKCAAEFLLMQYKDKKNNAEDLDEKLGKIIPEFRKQMYTFADYR 517
QY 111 DIFFGDISCRYIKDTSQTIKSKIGDQATTEKGDTHID-----NKKLQEMWTIHG 162
DB 518 DICLGTDIS---KKDTSKGV-----GKVKCIDDVFYKISINSIRKSMWETNG 564
QY 163 PKIMEGMLCALTNGLS-----ESEKKNIIQDYSYNKLNNAEKDDCCLKEFASKPOFLRW 216
DB 565 PVIMEGMLCALSDTSLNNVNPETHKKLTGNNNFEKVIIFGSDSTLTSFSERPQFLRW 624
QY 217 YVMSDFEGRERKLEKVEDVCIRAKDYEGCKNNKSNNSCVKV---CKEYENYITGKTK 273
DB 625 LTEMGENFECKGRK-EKVV--LAKCKDCVDGDCGCKNGCVACKQCKQYHSMIGIWD 681
QY 274 QYESQEGKENTERR--OKRPEYNSYKSKDASEYLK 306
DB 682 NYKKQGRYEVAKKIPLYKEDKDVKNKSDADYDLK 716

RESULT 11
Q95W83 PRELIMINARY; PRT; 2527 AA.
AC Q95W83;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Erythrocyte membrane protein 1 (Fragment).
GN VAR1.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxID=5833;
RN
  [1]
  SEQUENCE FROM N.A.
  RX MEDLINE=21442075; PubMed=11557894;
  RA Flick K., Scholander C., Chen Q., Fernandez V., Pouvelle B., Gysin J.,
  RA Wahlgren M.;
  RT "Role of nonimmune IgG bound to PFEMP1 in placental malaria."
  RL Science 293:2098-2100(2001).
  DR EMBL: AF36567; AAL12845.1; -.
  DR InterPro: IPR004258; PFEMP.
  DR Pfam: PF03011; PFEMP; 2.
  DR NON_TER 2527 2527
  SEQUENCE 2527 AA; 292137 MW; 59BE2766BF0425E4 CRC64;

Query Match 30.2%; Score 515.5; DB 5; Length 2527;
Best Local Similarity 36.4%; Pred. No. 5e-29;
Matches 120; Conservative 55; Mismatches 120; Indels 35; Gaps 10;

QY 5 SNEISGCPNK-ESYPMDCCKN-IDNSHSGACMPRRORCLVADLTGSGEIRKPEDILTK 58
DB 1253 NSDIGCIRKYKDGNEYPVWDCDTRNKIKTIGEGAVPMPRRORCLVADLTGQLD-QTDEK 1311
QY 59 ILRFINCAKETHEFAWHKRYKDVNAENELSGKIEGFRKQMYTFGDFRIFETGDISCR 105
DB 1312 LRDAFIKSAAEFLFLSHMKKKEKKENPTEEARSLDVEAQTLNDCIIEPEFRKQMYTF 1371
QY 106 FGDFRIFETGDISCRKYIKDTSQTIKSKIGDQATTEKGDTHIDNKKLQEMWTIHGPKT 165
DB 1372 FGDFRDLCLGCKDIGNDVIVNNNITTAFCNGAONPSDO-----DTDSORQVFWGTGKDI 1426

```

```

QY 166 WEGMICALNGLSESEKKNIIOD---YGVNKLNNNAEKDDCCLEKFAKSPQPIRLRVEMSD 222
DB 1427 WEGMICALNGLSESEKKNIIOD---YGVNKLNNNAEKDDCCLEKFAKSPQPIRLRVEMSD 1483
QY 223 EFCRERKKLEKVEDVC---IKAKDYEG-C-KNNKSNNSCVKCKEYENVTIGKKTQYSSQ 278
DB 1484 HFCRERKKLEKVEDVC---IKAKDYEG-C-KNNKSNNSCVKCKEYENVTIGKKTQYSSQ 1543
QY 279 EGFETERKKRKKPEVNSYSKDDASYLKD 308
DB 1544 KVFETDEKGYNDPDTIQTSTAYEYELKK 1573

RESULT 12
G904A2 PRELIMINARY: PRT: 1685 AA.
AC 0904A2: 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DR 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
RN Variant surface protein PfEMP1 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IT: MEDLINE=20144115; PubMed=10677532;
RX Planch J.D., Grais A.G., Kriek N., Hudson-Taylor D., Kyes S., Fagen T.,
RA Planch J.D., Baruch D.I., Newbold C.I., Miller L.H.,
RT Identification of a Plasmodium falciparum intercellular adhesion
RT molecule-1 binding domain: A parasite adhesion trait implicated in
RT cerebral malaria.
RN Proc. Natl. Acad. Sci. U.S.A. 97:1766-1771(2000).
DL EMBL; AF193424; AAF1980.1;
DR InterPro: IPR000454; Eub. AtPase_Csub.
DR InterPro: IPR004258; PFEMP.
DR Pfam: PF03011; PFEMP.
DR PROSITE: PS00605; ATPASE_C; 1.
FT NON TER 1685 1685
SQ SEQUENCE 1685 AA; 193671 MW; D1FD426666B0551E CRC64;

Query Match 26.0%; Score 444; DB 5; Length 1685;
Best Local Similarity 34.4%; Pred. No. 5.2e-24;
Matches 115; Conservative 46; Mismatches 111; Indels 42; Gaps 12;

```

```

RESULT 13
G900G6 PRELIMINARY: PRT: 2209 AA.
AC 0900G6: 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DR 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
RN Var. MAL4P2.56.
DE MAL4P2.56, VAR.
GN MAL4P2.56, VAR.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-3D7;
RA Devlin K., Bowman S., Churcher C., Harris B., Harris D., Lawson D.,
RA Quail M., Barrell B.,
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035475; CAB62897.1;
DR InterPro: IPR004258; PFEMP.
DR Pfam; PF03011; PFEMP.
SQ SEQUENCE 2209 AA; 250778 MW; 7A349F6FEFE9F7CB CRC64;

Query Match 25.6%; Score 437; DB 5; Length 2209;
Best Local Similarity 28.2%; Pred. No. 2.2e-23;
Matches 111; Conservative 51; Mismatches 114; Indels 118; Gaps 14;

QY 1 GNDGSEKNGSCNPK-----ESYDMDCC-----KKNIDNSHGACMPPRROKLCVRDL 47
DB 923 GNDGSD---ACGTIKYDKYGRKREKPPNMKCIAPSSDSNGSICVPPRRRLVYGGI 979
QY 48 TG-----GGERKKEDILTK-----FINCAKKTIPAHMKYKD-----81
DB 980 TKMASGNTVVSQAQTPQGGTSPSDNKLDAFIQSAIETFFLWKYMDKEIEKKQOO 1039
QY 82 -----NVNNEELK-----SGKIPEGFPRKQYTFGDPRIFFPGTDISSCYINDTSQ 129
DB 1040 KNLVANTSNVKGHEQKLEQSGIIPEDPRKQFYLGDRIIDLFKDIISGDKNMDTIE 1099
QY 130 TYSKLGDDATTEKGDTHIDNKKLOEWMTHNGPKTMEGMLALT-----174
DB 1100 KINGIILPKNGTSPSPAKNTP-----TEWMSONGEHIMNMICALTYDTNVSQDKPTQNE 1154
QY 175 -----NGLSESEKKNIILODYSN-----KLNNAEKDDCC-----LEKFAKSPQPI 214
DB 1155 KVEALMDENCKNPKNDYQSSVYTGEGAEGLQSTDSKDAARGKPTLDSITKRPYF 1214
QY 215 RMYVEMSDFCRERKKLEKVEDVCIKARDYEG-----CKNKSNN-----255
DB 1215 RYLEWGNONCKRERK---RLNDIKYECRGDENITRYSYGIGEDCKNNLPENPSTFKDLE 1271
QY 256 --SCVAVCKEYENVTIGKKTQYSSQEGKFNTEKR 287
DB 1272 YPTCARYCRFYKMKWMTKTEYKQEKIYVQOKK 1305

RESULT 14
G96108 PRELIMINARY: PRT: 1711 AA.
AC 096108: 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DR 01-DGC-2001 (TREMBLrel. 19, Last annotation update)
RN PFEMP1.
GN PFB0010W.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=99021743; PubMed=9804551;
RX Gardner M.J., Tetteilin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shallow S., Mason T., Yu K., Fujii C., Pederson J.,

```

RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
 RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
 RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
 RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
 falciparum";
 RL Science 282:1126-1132(1998).
 DR EMBL: AE001366; AAC71792.1; -
 DR InterPro: IPR004258; PFEMP.
 SO PFam: PF03011; PFEMP; 1.
 SO SEQUENCE 1711 AA; 196550 MW; 7BBD6C52742235B CRC64;

Query Match 25.4%; Score 433; DB 5; Length 1711;
 Best Local Similarity 33.4%; Pred. No. 3; 3e-23;
 Matches 104; Conservative 49; Mismatches 110; Indels 48; Gaps 15;

DB 20 WDC--KKNDNSHGA-CMPRRKQKIC-----VRDLTGGGELRKPEDILTKINCAKE 70
 856 WRCIANSQTTSKGDGACVPRPTEOLCLYLLKELSDPTQKG-----LEAFITKPADE 908
 71 TRFAWHKRYKKDNV-----AENELSGKIPGFRKQMYTFEGDFRIFGTDIS 119
 909 TYLLMCKYKEDKQNETASTFELDIDDPQTLNGSEIFEDFRQMFYTFGDRDLFLG---- 964
 DB 120 SCRTKDTSTKSLGDDQATTEKGDTHIDNKKL---OEWMTIHGPKIWEGMLCALIN 175
 965 --RYGNDLIDKYNNNI--TAVFQNGD-HIPNGQTROROEFGCTYKGDIMKMGALOE 1019
 QY 176 GLSESEKKNILDYSYNKLN-NAEKDDCCLEKFAKPOFLRWYEVMSDECRERKKLEDK 234
 DB 1020 A--GGKILLETNYNSNFTENGHLTGKTLNFAFSRSEFLRMWTEGDDQCRIRITQLOI 1076
 QY 235 VEDVICAKADYECCK-NKNSMNSCVKVEYENYITGKTKQYESQEGKFNTEROKKPEY 293
 DB 1077 LKRCM-VYQYNDKDKDKKCECTEACTYTKEMLNWMDNYKKQNGR-TEVKGISP-Y 1133
 QY 294 NYSKRDASEY 304
 DB 1134 KEDSDVSKSY 1144

RESULT 15

097312 PRELIMINARY; PRT; 2169 AA.
 ID 097312
 AC 097312
 DT 01-MAY-1999 (TEMBLrel. 10, Created)
 DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
 DT var (3D7-varT3-2).
 PF01120C, MAL3P7.55.
 Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3D7;
 RC MEDLINE=99376085; PubMed=10448855;
 RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
 RA Churcher C.M., Craig A., Davies R.M., Devlin K., Felwell T.,
 RA Gentles S., Gwilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
 RA Horrocks P., Jagels K., Jassal B., Kyes S., McLean J., Moule S.,
 RA Mungall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
 RA Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
 RA Whitehead S., Woodward J.R., Newbold C., Barrell B.G.;
 RT "The complete nucleotide sequence of chromosome 3 of Plasmodium
 falciparum";
 RL Nature 400:532-538(1999).
 DR EMBL: AL034559; CAB39061.1; -
 DR InterPro: IPR004258; PFEMP.
 DR Pfam: PF03011; PFEMP; 2.
 SO SEQUENCE 2169 AA; 246913 MW; 1D48ACB7AFAE3BE CRC64;

Query Match 23.9%; Score 407.5; DB 5; Length 2169;
 Best Local Similarity 25.8%; Pred. No. 3; 1e-21;

Matches 124; Conservative 52; Mismatches 116; Indels 189; Gaps 18;
 QY 10 GCNPKSYPPWDC-----KKNDNSH-----GACMPPR 38
 DB 850 GPEGKEKFPWKCITSSGSDTGTSTTKONDSBEGHRSKRKHTESSDSTTSSGSCVCPVR 909
 QY 39 ROKLCVRLDTQ--GGEIRK-----PED-----ILTKPINCARET 71
 DB 910 RRLVYVELTKMAEFARKSSTSPQPSGVANASASTSSPTDAPQLNDAFIQSAIET 969
 QY 72 HFAMHRYKKDNV-----NAENE-----LKSGLPEGFRKQMY 104
 DB 970 FFLMHRKYKAEETPDNKSPLNGGVAQFSGSYSGSESEKTPQOMLOSGLTPTDFLRQMF 1029
 QY 105 TFGDFRDIPEGDIISSCRYIKDTSQTKSLGQATTEK-----GDT-----HT 148
 DB 1030 TLGDYRDLGVYKEDYINALKASGDNPTNKLTIQIISEKIKVIEKSGDTPSPRTPPGQOP 1089
 QY 149 DNRKKLQEWMTINGPKIWEGMLCALT--NLSSESEK-----N 184
 DB 1090 SDNDP-KSMWKLAHPIYIMNGMIYALYDTFTASGEKKIEKDAVYKKLWDEANNRPKDN 1148
 QY 185 ILIDSYNKLNAEKD-----CCLKFAKPOFLRWYEVMSDE 223
 DB 1149 GOODTYEKEVEIEEDSGQKASTASQTPSPRASGENKPTLDSFVRKPTVFRYLEWGET 1208
 QY 224 FCRERKKLEKVDVC-----IKAK--DYEGCKNNKSNN-----SCV 258
 DB 1209 FCRERKKRLEKIVDCEVEENTGARGTTRKQKTSQGEDESKIVENDKLIFKOLEKPSCA 1268
 QY 259 KVCKEENYITGKTKQYESQEGKFNTEROKKPEYNSYSKK-----DASEYL 305
 DB 1269 TPCGLKRWIRKRKKDEYKQKSAVN---EQKITYENGNNKGGGNGVCGTLDENADFL 1324
 QY 306 K 306
 DB 1325 K 1325

Search completed: June 20, 2003, 15:05:23
 Job time : 50.1396 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 14:36:44 ; Search time 56.4583 Seconds
(without alignments)
828.417 Million cell updates/sec

Title: US-10-087-013-9

Perfect score: 1913
Sequence: 1 PCKMVKLISEQIEKNHNH.....AKHARDYKLTQENMCTNG 351

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Ched: 908470 seqs, 133250620 residues

Minimum number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_101002.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 1913 | 100.0 | 351 | 22 | AA62149 |
| 2 | 646 | 33.8 | 793 | 16 | AA70234 |
| 3 | 646 | 33.8 | 921 | 18 | AAW22480 |
| 4 | 646 | 33.8 | 921 | 21 | AAV77902 |
| 5 | 625.5 | 32.7 | 2703 | 16 | AAW70236 |
| 6 | 625.5 | 32.7 | 2710 | 18 | AAW22482 |
| 7 | 625.5 | 32.7 | 2710 | 18 | AAV77904 |
| 8 | 625.5 | 32.7 | 3060 | 18 | AAW22475 |
| 9 | 625.5 | 32.7 | 3060 | 21 | AAV77905 |
| 10 | 550.5 | 28.8 | 3542 | 22 | AA62142 |

| | | | | | | |
|----|-------|------|------|----|----------|---------------------|
| 11 | 548 | 28.6 | 407 | 22 | AA62148 | P. falciparum varC |
| 12 | 509.5 | 26.6 | 294 | 22 | AA62147 | P. falciparum varC |
| 13 | 444 | 23.2 | 308 | 22 | AA62151 | P. falciparum varC |
| 14 | 440 | 23.0 | 311 | 22 | AA62150 | P. falciparum varC |
| 15 | 412.5 | 20.8 | 2913 | 17 | AAW00384 | Plasmodium falcipa |
| 16 | 398.5 | 20.6 | 2197 | 21 | AA18352 | Plasmodium falcipa |
| 17 | 375 | 19.6 | 2228 | 20 | AAW3944 | Plasmodium falcipa |
| 18 | 374.5 | 19.6 | 2182 | 18 | AAW22476 | P. falciparum PTEM |
| 19 | 374.5 | 19.6 | 2182 | 21 | AAV77906 | Plasmodium var-1 P |
| 20 | 355 | 18.6 | 1726 | 17 | AAW00385 | Truncated Plasmodi |
| 21 | 353 | 18.5 | 749 | 16 | AA70233 | P. falciparum EBL- |
| 22 | 353 | 18.5 | 749 | 18 | AAW22479 | Plasmodium ebl-1. |
| 23 | 353 | 18.5 | 749 | 21 | AAV77901 | P. falciparum ebl- |
| 24 | 353 | 18.5 | 1086 | 23 | AAU76760 | Plasmodium falcipa |
| 25 | 333.5 | 17.4 | 1143 | 23 | AAU76759 | Plasmodium falcipa |
| 26 | 329 | 17.2 | 350 | 23 | ABW07656 | P. falciparum EBA- |
| 27 | 329 | 17.2 | 1435 | 16 | AA70232 | P. falciparum SABP |
| 28 | 329 | 17.2 | 1435 | 18 | AAW22477 | Silic acid bindin |
| 29 | 329 | 17.2 | 1435 | 21 | AAV77900 | P. falciparum SABP |
| 30 | 329 | 17.2 | 1604 | 16 | AA70105 | TNF-R-EBA 175 fusl |
| 31 | 329 | 17.2 | 1786 | 14 | AA41043 | CD4-EBA175 fusion |
| 32 | 312 | 16.3 | 1421 | 23 | AAU76764 | Plasmodium falcipa |
| 33 | 311 | 16.3 | 616 | 23 | AAW50533 | Unidentified amino |
| 34 | 303.5 | 15.9 | 1700 | 21 | AA18144 | Plasmodium falcipa |
| 35 | 284 | 14.8 | 324 | 21 | AAV77911 | Plasmodium DBL gen |
| 36 | 272.5 | 14.2 | 324 | 21 | AAV77915 | Plasmodium DBL gen |
| 37 | 269.5 | 14.1 | 1245 | 16 | AA70106 | TNF-R-Pl. vivax Du |
| 38 | 269 | 14.1 | 329 | 23 | ABW07654 | P. vivax PvDBP pro |
| 39 | 269 | 14.1 | 1028 | 14 | AA41044 | Plasmodium vivax Pl |
| 40 | 269 | 14.1 | 1061 | 16 | AA70231 | P. vivax DBP. Pl |
| 41 | 269 | 14.1 | 1115 | 12 | AA13457 | Duffy receptor. P |
| 42 | 269 | 14.1 | 1115 | 18 | AAW22478 | Duffy antigen bind |
| 43 | 269 | 14.1 | 1115 | 21 | AAV77899 | P. vivax DBP bind |
| 44 | 260 | 13.6 | 1807 | 22 | AAW5697 | Recombinant protei |
| 45 | 260 | 13.6 | 2028 | 22 | AAW5698 | Recombinant protei |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| AA62149 | |
| ID | AA62149 standard; peptide: 351 AA. |
| AC | AA62149; |
| XX | |
| DT | 29-MAY-2001 (first entry) |
| XX | |
| DE | P. falciparum varCSA polypeptide AttresDBL3-gamma. |
| XX | |
| KW | FCR3, varCSA protein; chondroitin sulfate A; CSA; var gene; PFEWPI; |
| KW | erythrocyte membrane protein 1; parasitized red blood cell; PFRBC; |
| KW | malaria; protozoa; AttresDBL3-gamma. |
| XX | |
| OS | Plasmodium falciparum. |
| XX | |
| PN | W0200116326-A2. |
| XX | |
| PD | 08-MAR-2001. |
| XX | |
| PE | 01-SEP-2000; 2000WO-US24195. |
| XX | |
| PR | 01-SEP-1999; 99US-0152023. |
| XX | |
| PA | (USSH) US DEPT HEALTH & HUMAN SERVICES. |
| XX | |
| PI | Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C; |
| XX | Gysin J, Pouvelle B, Fujii N, Smith J; |
| DR | WPI: 2001-235109/24. |
| XX | |
| PT | Novel FCR3, varCSA protein, useful for modulating parasitized red blood |
| PT | cell binding, sequestration and onset of maternal malaria - |

XX Claim 54; Page 73-74; 78pp; English.

PS The invention relates to a P. falciparum FC3, varCSA protein, that is

XX capable of binding to Chondroitin sulfate A (CSA). The var gene and the

CC corresponding P. falciparum erythrocyte membrane protein 1 (PEEMP1)

CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The

CC protein and the encoding gene are useful for treating and preventing

CC maternal malaria or in a patient afflicted at a risk for contracting

CC maternal malaria or in a patient afflicted with maternal malaria. The

CC present sequence represents a P. falciparum varCSA polypeptide

CC AddressBuj-panama.

XX

SQ Sequence 351 AA;

Query Match 100.0%; Score 1913; DB 22; Length 351;

Best Local Similarity 100.0%; Pred. No. 3.6e-167;

Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PCKMVKLISEQIEKNNIHNCKKTEDAKMKCENTKLGEDEGVCMPRRONLCVHYLTKLN 60

1 PCKMVKLISEQIEKNNIHNCKKTEDAKMKCENTKLGEDEGVCMPRRONLCVHYLTKLN 60

Db 61 DSKSEEDLRFAFIKSAAEFTLLROYNSKNVEDDKILHRDMLPPEFFRSMEYTFGDIRD 120

QY 61 DSKSEEDLRFAFIKSAAEFTLLROYNSKNVEDDKILHRDMLPPEFFRSMEYTFGDIRD 120

Db 61 DSKSEEDLRFAFIKSAAEFTLLROYNSKNVEDDKILHRDMLPPEFFRSMEYTFGDIRD 120

QY 121 ICIDTDISEKIDHDVTTAKKITTAVFQKIGSKTNTKKVLEREGMKKEVGLSTWKMCLC 180

121 ICIDTDISEKIDHDVTTAKKITTAVFQKIGSKTNTKKVLEREGMKKEVGLSTWKMCLC 180

Db 121 ICIDTDISEKIDHDVTTAKKITTAVFQKIGSKTNTKKVLEREGMKKEVGLSTWKMCLC 180

QY 181 ALSYNETKMKDESVRTYLMKKYIKNNDIKEYLEEFASRPPLRWVTEMEDEFFKNNRKE 240

181 ALSYNETKMKDESVRTYLMKKYIKNNDIKEYLEEFASRPPLRWVTEMEDEFFKNNRKE 240

Db 181 ALSYNETKMKDESVRTYLMKKYIKNNDIKEYLEEFASRPPLRWVTEMEDEFFKNNRKE 240

QY 241 LVSLAKKDCSCTLRNNGTSKCTCDNENCGACTQCEKRYKMKERKKHYSQKKRQLY 300

241 LVSLAKKDCSCTLRNNGTSKCTCDNENCGACTQCEKRYKMKERKKHYSQKKRQLY 300

Db 241 LVSLAKKDCSCTLRNNGTSKCTCDNENCGACTQCEKRYKMKERKKHYSQKKRQLY 300

QY 301 KNSATYNGLAVEANSEYTKNDPEVTEANSKARDYLTQLENNMCTNG 351

301 KNSATYNGLAVEANSEYTKNDPEVTEANSKARDYLTQLENNMCTNG 351

Db 301 KNSATYNGLAVEANSEYTKNDPEVTEANSKARDYLTQLENNMCTNG 351

RESULT 2

AAAR70234

ID AAR70234 standard; Protein; 793 AA.

XX

AC AAR70234;

XX

XX 22-SEP-1995 (first entry)

XX

DE P. falciparum E31a.

XX

KW Erythrocyte binding ligand; E31a; binding domain; malaria; therapy;

XX vaccine.

XX

OS Plasmodium falciparum.

XX

PN W09507353-A.

XX

PD 16-MAR-1995.

XX

XX 07-SEP-1994; 94W0-US10230.

XX

PF 10-SEP-1993; 93US-0119677.

XX

PR

XX (USSH) US REPT HEALTH & HUMAN SERVICES.

XX

PA Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;

XX PI Wellens TE;

XX WPI; 1995-123427/16.

DR N-PSDB; AA083527.

XX

XX New erythrocyte binding domain polypeptide(s) - isolated from

PT Plasmodium binding proteins, used in diagnosis, treatment and

PT prevention of malaria

XX

PS Disclosure; Page 51-52; 81pp; English.

XX

CC Erythrocyte binding ligand (EBL) family genes were cloned from

CC P. falciparum chromosome 7 subsegment libraries constructed during

CC genetic studies of the chloroquine resistance locus. The 4 genes,

CC EBL-e1 (AA083526), E31a (AA083527), EBL-e2 (AA083528) and ProJ3

CC (AA083529), encode the proteins given in AAR70233-36, respectively. The

CC binding domains of such proteins can be expressed e.g. in E. coli.

CC yeast, mammalian, insect and in vacchia virus and adenovirus-infected

CC cells, and provide protection against P. falciparum.

XX

SQ Sequence 793 AA;

Query Match 33.8%; Score 646; DB 16; Length 793;

Best Local Similarity 41.0%; Pred. No. 2.9e-50;

Matches 150; Conservative 54; Mismatches 112; Indels 50; Gaps 16;

QY 2 CKMVKLISEQIEKNNIHNCKKTED-AKMKCENT--KLGEDEGVCMPRRONLCVHYLTK 58

396 CEIYAEMLKDKNGRTYVGEYRKREYSEWTCDESKIRMG-HGACIPPRQKLCIHYLEK 454

Db 59 LNDSEKEDLRFAFIKSAAEFTLLROY--NSKNVED-DKILHRDMLPPEFFRSMEYTF 114

455 IMNTNTE--LKYAFIKCAAEFTLLWQNYKKDKNGNAEDLDEKLGITIPEDFRQMFYR 512

QY 115 FGDRDLCIDTDI-SEKIDHDVTTAKKITTAVFQKIGSKTNTKKVLEREGMKKEVGLS 173

513 FADYRDLCLDTDISSKKDTSKYGVKACNIDVYFKISN-----SIRYKSWETNGPY 566

Db 174 IMKGMICALSYNT-----ET-KKMDGVRTYLMKKYIKNNDIKEYLEEFASRPPLRW 225

567 IMEGMICALSYDPSLNNVNPEHKKLLEGNNNE-EKVIYF-GSDSFTLLSFSPQFLRW 624

QY 226 VTEWGEDFVNNRKKELVSLKKDCSCTLRNNGTSKCTCDNENCGACTQCEKRYKMKER 285

625 LTFWGENFECKQKREYKVLAKCKDCDVGDDG---KC--NGKCVACKDCQKQYHSMWIGI 678

Db 286 WKKHYSQKKKPOLYKNSATYNGLAVEANSEYTKNDPEVTEANSKARDYLTQLENN 345

679 WIDNYKKQKRGRTYEVK-----KILPYKEDKRYKNSD---ARDYLTQLOLYN 721

QY 346 MICTNG 351

Db 722 MKCVNG 727

RESULT 3

AAW22480

ID AAW22480 standard; Protein; 921 AA.

XX

AC AAW22480;

XX

XX 07-OCT-1997 (first entry)

XX

DE Plasmodium E31a.

XX

XX DBL gene family; SABP; sialic acid binding protein; vaccine; therapy;

KW Duffy binding like gene; Duffy antigen binding protein; erythrocyte;

KW DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;

XX Plasmodium.

XX

OS Plasmodium falciparum.

XX

PN W09640766-A2.

XX

PD 19-DEC-1996.

| | | | |
|----------|-------------|---|------|
| Db | 1462 | YRDIQCLNTDISKK--QNDVAKAKKDKIGKFFSNDGSKSPSG--LSRQEMWKTNPGEITKG | 1516 |
| Oy | 178 | MCAALS-VTEFFKKNDEGVRTYLMKYYKKNNDIKEEYLEEFPASRPFLRWYEMGEDPYKN | 236 |
| Db | 1517 | MCALIKRYTDTIDPNRKIKNDYSYDKRVQNSGNSPLEEPAKPGFLKMLIEMGEPCAE | 1576 |
| Oy | 237 | RKKEVLSEKKKCDSCILRNNGTSNKTCDNENCGACKQCEKYYKMMERKKKHSSOKK | 296 |
| Db | 1577 | RKKKEIITIDAC-----NEINSTQOCNDAKH--RCNACRAYQRYEYENKKEEFGSQTN | 1628 |
| Oy | 297 | FOLYKNSATYINGLAVKEANSSETYKNDPEVTANSAKHARDYLTQLENICT | 349 |
| Db | 1629 | FVLKANVQPOD-----PEYKGYEKYDGVPIQGN-----EYLQKQCDNNKS | 1670 |
| RESULT 6 | | | |
| ID | AAW22482 | standard; Protein; 2710 AA. | |
| AC | AAW22482; | | |
| XX | | | |
| XX | 07-OCT-1997 | (first entry) | |
| DT | | | |
| XX | | | |
| De | | Plasmodium Proj3. | |
| KW | | DBL gene family; SABB; sialic acid binding protein; vaccine; therapy; | |
| KM | | Duffy binding like gene; Duffy antigen binding protein; erythrocyte; | |
| KW | | DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response; | |
| KW | | Plasmodium. | |
| OS | | Plasmodium falciparum. | |
| XX | | | |
| PN | | WO9640766-A2. | |
| XX | | | |
| PD | | 19-DEC-1996. | |
| XX | | | |
| PF | | 07-JUN-1996; 96WO-US09508. | |
| PR | | 07-JUN-1995; 95US-0487826. | |
| XX | | | |
| PA | | (USSH) US DEPT HEALTH & HUMAN SERVICES. | |
| XX | | | |
| PI | | Chitnis C, Miller LH, Peterson DS, Sim KL, Su X; | |
| PI | | Wellens TE; | |
| XX | | | |
| DR | | WPI: 1997-052231/05. | |
| XX | | N-PSDB: AAT72897. | |
| PT | | New malaria vaccines - contains cysteine-rich DBL family protein | |
| PT | | binding domains homologous domains of the Duffy and sialic acid | |
| XX | | binding proteins | |
| PS | | Disclosure; Page 50-56; 96pp; English. | |
| XX | | | |
| XX | | | |
| CC | | This sequence represents Proj3 of Plasmodium. Proj3 belongs to | |
| CC | | the duffy binding like (DBL) family of genes which have homology to the | |
| CC | | Duffy antigen binding protein (DABP) and sialic acid binding protein | |
| CC | | (SABP) conserved regions (see AAT72889 and AAT72888 respectively). The | |
| CC | | var family of genes modulate cytoadherence and antigenic variation of | |
| CC | | Plasmodium infected erythrocytes. SABP and the Duffy antigen binding | |
| CC | | protein (DABP) are soluble proteins that appear in the culture | |
| CC | | supernatant after infected erythrocytes release merozoites. DABP and | |
| CC | | SABP mediate the binding of merozoites and schizonts to the erythrocyte | |
| CC | | surface. These proteins are necessary for erythrocyte invasion by the | |
| CC | | parasite. This sequence can be used in the compositions of the | |
| CC | | invention. The compositions are for the treatment and prevention of | |
| CC | | malaria, and comprise either a nucleotide sequence or encoded polypeptide | |
| CC | | of the var-1, var-2, var-3 or var-7 genes of the DBL gene family, a | |
| CC | | family of genes having homology with conserved regions of DABP and SABP. | |
| CC | | The compositions are used for the treatment and prevention of malaria. | |
| CC | | They are also used in the preparation of vaccines for inducing a | |
| CC | | protective immune response in a mammal to Plasmodium merozoites | |
| CC | | (especially Plasmodium falciparum or Plasmodium vivax). | |

XX Sequence 2710 AA: 32.7% Score 625.5; DB 18; Length 2710;
 Query Match Best Local Similarity 38.8%; Pred. No. 1.3e-47;
 Matches 137; Conservative 61; Mismatches 126; Indels 29; Gaps 9;

2 CKWVKLISQIEKNNHNCCKTED-AKKCENTKLGEDGYCMPPRONLCVHYL---T 57
 1349 CKIVEKILBGNKGRVTGECNPKRESYPDMCKNNIDISHDGACMPPRKCLCYTAHES 1408

QY 58 KLNDKSKEDLRFAFKSAAEFTLLROYNSKNVEDDKILHDMIPPEFRSMFTFGD 117
 1409 QTEINIKTDNLKDAFIKTAAEFTLSWQYKSKNDSAKILDRGLIPSOFLRSMYTFGD 1468

QY 118 YRDICLDTISEKIAHDVTTAKKITAFAVKIGSTGKTTNGKVLERBGMKKEYGLSTWG 177
 1469 YRDICLNTDISKR--ONDAKAKADKIGKFFSKDGSKSPSG---LSROEMWKTGPEIWK 1523

Db 178 MLCALSL-YNTERKKMDGVTYLMKYIYKNNDIKEYLEEFASRPFLRWTEWGEDFVK 236
 1524 MLCALFKRYVTDNKRRIKNDYSYDKVNSQNSNPISLEFPAKPOFLRMWIEGEEFCAL 1583

QY 237 RKRELVSLEKKCDSCFLRNNGTSNKTCDNENCGACKTQCEKTKMMERKKHYSSOKK 296
 1584 ROKKENIIRKDAC-----NEINSTQOCNDAKH--RCNOACRAVOEYVENKKKEFSQGTNN 1635

QY 297 FOLYKNSATYNNGLAVKEANSETYKNDPEYTEANSAKHARDYKLTQLENNICT 349
 1636 FVLKANVQPOD-----PEYKGYEYKDGVPQIQGN-----EYLQKCDNNKCS 1677

Db 1636 FVLKANVQPOD-----PEYKGYEYKDGVPQIQGN-----EYLQKCDNNKCS 1677

RESULT 7
 AAY77904
 ID AAY77904 standard; protein; 2710 AA.
 AC AAY77904;
 DT 13-JUN-2000 (first entry)

DE P. falciparum Proj3 binding domain polypeptide.
 XX DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
 KW DABP; Sialic Acid Binding Protein; SABP; malaria; vaccine; immunisation;
 KM protozoicide; Proj3.
 CC Plasmodium falciparum.

US5993827-A.
 PD 30-NOV-1999.
 PF 07-JUN-1995; 950S-0487826.
 PR 10-SEP-1993; 930S-0119677.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Sim KL, Chittis C, Peterson DS, Su X, Wellens TE, Miller LH;
 DR WPI; 2000-194198/17.
 DR N-PSDB; AA298286.

PT Isolated protein binding domains from Plasmodium vivax and Plasmodium
 PT falciparum erythrocyte binding proteins useful for vaccinating against
 PT malaria -
 PS disclosure; Columns 79-92; 93pp; English.

XX The invention relates to ebl-1 polypeptides that are encoded by the DBL
 CC (Duffy-binding like) gene family. The ebl-1 proteins are substantially
 CC identical to the Duffy Antigen Binding Protein (DABP) and Sialic
 CC Binding Protein (SABP), which are soluble proteins that appear in the

CC culture supernatant after erythrocytes infected with malaria release
 CC merozoites. Immunochemical studies indicate that DABP and SABP are the
 CC respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy
 CC and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be
 CC used to vaccinate against malaria, especially caused by P. falciparum.
 CC Immunization with the polypeptide provides effective protection against
 CC malaria. The present sequence represents the Proj3 binding domain
 CC polypeptide.

QY 2 CKWVKLISQIEKNNHNCCKTED-AKKCENTKLGEDGYCMPPRONLCVHYL---T 57
 1349 CKIVEKILBGNKGRVTGECNPKRESYPDMCKNNIDISHDGACMPPRKCLCYTAHES 1408

QY 58 KLNDKSKEDLRFAFKSAAEFTLLROYNSKNVEDDKILHDMIPPEFRSMFTFGD 117
 1409 QTEINIKTDNLKDAFIKTAAEFTLSWQYKSKNDSAKILDRGLIPSOFLRSMYTFGD 1468

QY 118 YRDICLDTISEKIAHDVTTAKKITAFAVKIGSTGKTTNGKVLERBGMKKEYGLSTWG 177
 1469 YRDICLNTDISKR--ONDAKAKADKIGKFFSKDGSKSPSG---LSROEMWKTGPEIWK 1523

QY 178 MLCALSL-YNTERKKMDGVTYLMKYIYKNNDIKEYLEEFASRPFLRWTEWGEDFVK 236
 1524 MLCALFKRYVTDNKRRIKNDYSYDKVNSQNSNPISLEFPAKPOFLRMWIEGEEFCAL 1583

QY 237 RKRELVSLEKKCDSCFLRNNGTSNKTCDNENCGACKTQCEKTKMMERKKHYSSOKK 296
 1584 ROKKENIIRKDAC-----NEINSTQOCNDAKH--RCNOACRAVOEYVENKKKEFSQGTNN 1635

QY 297 FOLYKNSATYNNGLAVKEANSETYKNDPEYTEANSAKHARDYKLTQLENNICT 349
 1636 FVLKANVQPOD-----PEYKGYEYKDGVPQIQGN-----EYLQKCDNNKCS 1677

Db 1636 FVLKANVQPOD-----PEYKGYEYKDGVPQIQGN-----EYLQKCDNNKCS 1677

RESULT 8
 AAW22475
 ID AAW22475 standard; protein; 3060 AA.
 AC AAW22475;
 DT 12-SEP-1997 (first entry)

DE Plasmodium var-7.
 XX DBL gene family; SABP; sialic acid binding protein; vaccine; therapy;
 KW Duffy binding like gene; Duffy antigen binding protein; erythrocyte;
 KW DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;
 KM Plasmodium.
 XX Plasmodium vivax.
 OS Plasmodium falciparum.
 PN WO9640766-A2.
 PD 19-DEC-1996.
 PF 07-JUN-1996; 96WO-US09508.
 PR 07-JUN-1995; 950S-0487826.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Chittis C, Miller LH, Peterson DS, Sim KL, Su X;
 DR WPI; 1997-052231/05.
 DR N-PSDB; AAT72882.

XX
AC
AA00214Z
DT
29-MAY-2001 (first entry)

DE P. falciparum FCRI3.varCSA protein.
 XX FCRI3.varCSA protein; chondroitin sulfate A (CSA); var gene; PFEMP1;
 XX erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
 KM malaria; protozoacide.
 XX Plasmodium falciparum.
 OS WO200116326-A2.
 XX
 XX
 PD 08-MAR-2001.
 XX
 XX 01-SEP-2000; 2000WO-US24195.
 PF
 XX 01-SEP-1999; 99US-0152023.
 PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;
 XX Gysin J, Pouvelle B, Fujii N, Smith J;
 DR WPI: 2001-235109/24.
 XX N-PSDB; AAF57301.
 DR
 XX Novel FCRI3.varCSA protein, useful for modulating parasitized red blood
 PT cell binding, sequestration and onset of maternal malaria -
 XX
 PS Claim 12: Page 63-71; 78pp; English.
 XX
 XX The invention relates to a P. falciparum FCRI3.varCSA protein, that is
 CC capable of binding to chondroitin sulfate A (CSA). The var gene and the
 CC corresponding P.falciparum erythrocyte membrane protein 1 (PFEMP1)
 CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
 CC protein and the encoding gene are useful for treating and preventing
 CC maternal malaria in a patient identified at a risk for contracting
 CC maternal malaria or in a patient afflicted with maternal malaria. The
 CC present sequence represents the P. falciparum FCRI3.varCSA protein.
 CC
 SQ Sequence 3542 AA:
 Query Match 28.8%; Score 550.5; DB 22; Length 3542;
 Best Local Similarity 39.0%; Pred. No. 1.5e-40;
 Matches 133; Conservative 48; Mismatches 115; Indels 45; Gaps 14;
 QY 2 CKWVKLSLQJLKNHNC--KKTEDA--KMKCENTKGEDEGVCMPPRQNLGVHYLT 57
 1261 CKIVNDILKENDGCKKQYEDCHPKKNSNGIPDMQGINLYEDPRVCPPRQKLCVHFLA 1320
 58 KLNDGSK---EEDLREAFITKSAAEFTLLRQYNSKNVED--DKILHRDMPPEFFSRM 111
 1321 NDREIKKLSQVNLKFAFIKSAAEFTFSWYKSKDGEENELDKELKEGKIPAPFLRSW 1380
 QY 112 FYTFGGYRDLCDTIDSEKADHDVTTAKKITAIVQKISKTTNGKKVLEBGMKEYG 171
 1381 FYTFGGYRDLFTDIDSK--GHGEGSKLEQIDSLFKNGDOKSPNK--TROEMWTEHS 1435
 QY 172 LSIWKMICALSYNTEFKMKDEGVRTYLMKYIKN--NDIKREYLEEFASRPFLRWATE 228
 1436 HEIWEAMLCAL-VKIAKKDD-----FTENYGINNVKFSKSTLLEFARPPFLRLUTE 1489
 QY 229 WGEDFYVNRKKELVLSKKKCDSCFLRNNGTSNKTCDNENCGA-CTQCEKRYKWRMRK 287
 1490 WYDYCYTRQKLYDQEKCS-----NDQLKCDTECKNKKCEDYKYNKK-K 1535
 QY 288 KHYSSOKKKFQLYKNSATYNN---GLAVKE---ANSETYKN 322
 1536 KEMIPQDKYYKDERDKKRFDRQHIGVAVTDYTGTMATDLYN 1576
 DB
 RESULT 11
 AAB62148
 ID AAB62148 standard; peptide; 407 AA.
 XX

AC AAB62148;
 XX
 XX 29-MAY-2001 (first entry)
 DT
 XX P. falciparum varCSA polypeptide A4 DBL4-gamma.
 DE
 XX FCRI3.varCSA protein; chondroitin sulfate A (CSA); var gene; PFEMP1;
 XX erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
 KM malaria; protozoacide; A4 DBL4-gamma.
 XX Plasmodium falciparum.
 OS WO200116326-A2.
 XX
 XX
 PD 08-MAR-2001.
 XX
 XX 01-SEP-2000; 2000WO-US24195.
 PF
 XX 01-SEP-1999; 99US-0152023.
 PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;
 XX Gysin J, Pouvelle B, Fujii N, Smith J;
 DR WPI: 2001-235109/24.
 XX
 XX Novel FCRI3.varCSA protein, useful for modulating parasitized red blood
 PT cell binding, sequestration and onset of maternal malaria -
 XX
 PS Disclosure: Page 72-73; 78pp; English.
 XX
 XX The invention relates to a P. falciparum FCRI3.varCSA protein, that is
 CC capable of binding to chondroitin sulfate A (CSA). The var gene and the
 CC corresponding P.falciparum erythrocyte membrane protein 1 (PFEMP1)
 CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
 CC protein and the encoding gene are useful for treating and preventing
 CC maternal malaria in a patient identified at a risk for contracting
 CC maternal malaria or in a patient afflicted with maternal malaria. The
 CC present sequence represents a P. falciparum varCSA polypeptide
 CC A4 DBL4-gamma.
 CC
 SQ Sequence 407 AA:
 Query Match 28.6%; Score 548; DB 22; Length 407;
 Best Local Similarity 34.6%; Pred. No. 1.1e-41;
 Matches 126; Conservative 52; Mismatches 128; Indels 58; Gaps 12;
 QY 10 SEQIER-NNIHNCKTDEDKMKCENTKLGEDGVCMPPRQNLGVHYLTKLNDSEEDL 68
 6 TKKIEGNTKRYPTKNDYPMWNTDQVINEBSCMPPRQKLCINHLHLSKATETEL 65
 QY 69 REAFITSAAEFTLLRQYNSK-----NVEDD-----KILHRDMPPEFFSMRY 113
 66 RKAFICCAAEFTPLMDKYEDKDEKTEGGISDPDDPQKLLGCTIPEDFFKQMFY 125
 QY 114 TFGDYRDLCDTIDSEKADHDVTTA-KKITAIVQKISKTTNGKKVLEBGMKEYG 172
 126 TYGDYRDLFTDIDSK--GHGEGSKLGRKIDSLFKNGDOKSPSGKTPTE--WMNDYGP 179
 QY 173 SIWKMICALSYNTEFKMKDEGVRTYLMKYIKN--DIKEYLEEFASRPFLRWATE 229
 180 DIWKMVCGLSHH-----INGKKEQRLKRLTDNNYTKLSSLEDPASRPQFLRWIEW 234
 QY 230 GEDFYVNRKKELVLSKKCD--SCLLRNNGTSNKTCDNENCGACATQCEKRYKWRMRK 287
 235 GDQFCRERYVAKINQLTGCMNEYEGSGSENGKKE-----ACKNACAVKSWKDKM 284
 QY 288 KHYSSOKKKFQLYKNSATYNNGLAVKEANSETYKNPDEVTEANSKARHRYLKTOLNMI 347
 285 DQYEQTAQFD-----KDKKDKKFDGTSAEVVAVSSVHEYLDQELKN-L 329
 QY 348 CTNG 351
 DB

Db 330 CTG 333

RESULT 12

AAB62147 ID AAB62147 standard; peptide; 294 AA.

XX AAB62147;
AC
XX
DT 29-MAY-2001 (first entry)
XX
XX P. falciparum varCSA polypeptide R29DBL2-gamma.
DE
XX FCR3 varCSA protein; chondroitin sulfate A; CSA; var gene; PFEMP1;
KW erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
KW malaria; protozoacide; R29DBL2-gamma.
XX
OS Plasmodium falciparum.
XX WO200116326-A2.

08-MAR-2001.

XX 01-SEP-2000; 2000WO-US24195.

XX 01-SEP-1999; 99US-0152023.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;
PI Gysin J, Pouvelle B, Fujii N, Smith J;
XX WPI: 2001-235109/24.

XX Novel FCR3 varCSA protein, useful for modulating parasitized red blood
PT cell binding, sequestration and onset of maternal malaria -
XX

PS Disclosure; Page 72; 78pp; English.

XX The invention relates to a P. falciparum FCR3 varCSA protein, that is
CC capable of binding to chondroitin sulfate A (CSA). The var gene and the
CC corresponding P. falciparum erythrocyte membrane protein 1 (PFEMP1)
CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
CC protein and the encoding gene are useful for treating and preventing
CC maternal malaria in a patient afflicted at a risk for contracting
CC maternal malaria or in a patient afflicted with maternal malaria. The
CC present sequence represents a P. falciparum varCSA polypeptide
R29DBL2-gamma.

Sequence 294 AA;

Query Match 26.6%; Score 509.5; DB 22; Length 294;

Best Local Similarity 37.2%; Pred. No. 2.4e-38;
Matches 121; Conservative 46; Mismatches 101; Indels 57; Gaps 13;

QY 4 MVOKLISEQLEKNNHCKKTEDAKKKE-NTKLGDEGVCMPPRRONLCVHYLTKLND 62
Db 1 VLOOKSGSDNCNAKRNKNE--WODKNTFYDNGGVCMPPRRKSICHLNLTLEQT 57
QY 63 SKEDLEAFIKSAAEFTFL-ROYNSKNVEDDKILHRDMPPEFFRSMEFTGDRDI 121
Db 58 KKKQLAEAFIKCAKETNLMMDKYKNDK-EAEELLKKGKIPEDFPRIMYTTGDRDF 116
QY 122 CLDPTDISEKLAHDVYTTAKKKTIVFQKIGSKTTNGKVLEREGMMKEYGLSTIKGMICA 181
Db 117 CLENDMKR---DVDYKKNINRVF--NNSSKRGFKKIDPENWMNENGQIINGMICA 168
QY 182 LSYVTEKKMDGVRTILMKYIKNNDIKE-----LEEFASPPFLRW 225
Db 169 L-TIADTKRDS-----TKNDKYKTEKVTILAKRGDSGMILSEPAKKPKFLRW 215
QY 226 VTEWGEDFVKNRKKEVLSKKKCDSTLRNNGTSNKTCDNENCGACKTQCEKYYKWMER 285

Db 216 FVEWYDYCKEROKYLEVASTCKST----DGGOLK-CD-----RCGNKNCDEKXYMRK 265

QY 286 WKHHYSSQKKRFFOLYKNSATYNNGL 310

Db 266 KKEEMNLQDK-----YKDKRENKGI 286

RESULT 13

AAB62151 ID AAB62151 standard; peptide; 308 AA.

XX AAB62151;
AC
XX
DT 29-MAY-2001 (first entry)
XX
XX P. falciparum varCSA polypeptide ItG2-CS2 DBL2.
DE
XX FCR3 varCSA protein; chondroitin sulfate A; CSA; var gene; PFEMP1;
KW erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
KW malaria; protozoacide; ItG2-CS2 DBL2..

XX Plasmodium falciparum.

XX WO200116326-A2.

XX 08-MAR-2001.

XX 01-SEP-2000; 2000WO-US24195.

XX 01-SEP-1999; 99US-0152023.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;
PI Gysin J, Pouvelle B, Fujii N, Smith J;
XX WPI: 2001-235109/24.

XX Novel FCR3 varCSA protein, useful for modulating parasitized red blood
PT cell binding, sequestration and onset of maternal malaria -
XX

PS Claim 54; Page 75-76; 78pp; English.

XX The invention relates to a P. falciparum FCR3 varCSA protein, that is
CC capable of binding to chondroitin sulfate A (CSA). The var gene and the
CC corresponding P. falciparum erythrocyte membrane protein 1 (PFEMP1)
CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
CC protein and the encoding gene are useful for treating and preventing
CC maternal malaria in a patient afflicted at a risk for contracting
CC maternal malaria or in a patient afflicted with maternal malaria. The
CC present sequence represents a P. falciparum varCSA polypeptide
ItG2-CS2 DBL2.

Sequence 308 AA;

Query Match 23.2%; Score 444; DB 22; Length 308;

Best Local Similarity 34.4%; Pred. No. 2.7e-32;
Matches 115; Conservative 46; Mismatches 131; Indels 42; Gaps 12;

QY 16 NNINCKRTD-AKKKCENTKLGDEGVCMPPRRONLCVHYLTKLNDSEEDLREAFIK 74
Db 6 NEISGCKPRKESYPMDCNKINDSHSGACMPRRQKLCVADLGGGTRRKPEDILTFIN 65
QY 75 SAAAEFTLLROYNSKNVEDDKILHRDMPPEFFRSMEFTGDRDICLPTDISEKLAHD 134
Db 66 CAAKETFFAHMKYKKNVNAENELKSGKIDEGFRKQWYTFGDRDFFGTDISSCRYIK 125
QY 135 DVT-TAKKKT--TAVFQKIGSKTTNGKVLEREGMMKEYGLSTIKGMICALSYN-PTTK 190
Db 126 DTSQITISKLGDAQTTTKGQTHIDNKKID--WVTHQPKTWEGLCALTLNGLSSEK 182
QY 191 MDEGVRTYLMKYIK--NNDIKE--YLEEFASPPFLRWVTEWGEDFVKNRKKEVLSKK 246


```

Db      183 -----KNILQDYSYKLNNAEKDDCLLEFASKPOFLRMVYEMSDPEFCERRK-----LED 233
QY      247 KCDSCITLRNNGTSNKTCDNENCGACKTQCEKYKMMERMKHHYSQKKFOLYKNSATY 306
      234 KVEDVICR--AADEYECCKNNKSNNSCVKRYCKEYENITGKKTQYESQEGKF-----282
QY      307 NNGLAVKEANSEYKNDPEVTEANSKAKHARDYLK 340
Db      283 -----NTEKROKKPEY-NSYSKKDASEYIK 306

RESULT 14
AAB62150
ID AAB62150 standard; peptide; 311 AA.
AC AAB62150;
XX
XX 29-MAY-2001 (first entry)
XX
XX P. falciparum varCSA polypeptide FC3 var3DBL-gamma.
XX
XX FC3, varCSA protein; chondroitin sulfate A; CSA; var gene; PFEPM1;
XX erythrocyte membrane protein 1; parasitized red blood cell; PBC;
XX malaria; protozoacide; FC3 var3DBL-gamma.
XX
XX Plasmodium falciparum.
XX
XX WO200116326-A2.
XX
XX 08-MAR-2001.
XX
XX 01-SEP-2000; 2000MO-US24195.
XX
XX 01-SEP-1999; 99US-0152023.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;
XX Gysin J, Pouvelle B, Fujil N, Smith J;
XX WPI; 2001-235109/24.
XX
XX Novel FC3, varCSA protein, useful for modulating parasitized red blood
XX cell binding, sequestration and onset of maternal malaria -
XX
XX Disclosure; Page 74-75; 78pp; English.
XX
XX The invention relates to a P. falciparum FC3, varCSA protein, that is
XX capable of binding to chondroitin sulfate A (CSA). The var gene and the
XX corresponding P.falciparum erythrocyte membrane protein 1 (PFEPM1)
XX modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
XX protein and the encoding gene are useful for treating and preventing
XX maternal malaria in a patient afflicted at a risk for contracting
XX maternal malaria or in a patient afflicted with maternal malaria. The
XX present sequence represents a P. falciparum varCSA polypeptide
XX FC3 var3DBL-gamma.
XX
XX
XX Sequence 311 AA;
XX
XX Query Match 23.0%; Score 440; DB 22; Length 311;
XX Best Local Similarity 31.9%; Pred. No. 6,3e-32;
XX Matches 109; Conservative 52; Mismatches 133; Indels 48; Gaps 9;
QY      14 EKNINHC-----KTEDAKKCENTKAGEDEGVCMPPRRNLVCHVLTLNDSKEED 67
      1 EPTDIDGCMQKYKAGDKYPMDCNSQIHTTHNGACMPPRKLCVSGLTTRIKAIEX 60
QY      68 LREAFIKSAAEFELLROYNSKNVEDDKLHRDMIPPEFERSMEYTFGDIRICLDTPI 127
      61 IRTEFIKSAALIEHFMDRKEDNGAELEKNGINPEGFKRMVITFGDIRIFGRDI 120
QY      128 S-----EKIADHDVTTAKKKTAVFOKISKTNGKVLREEGMKREYGLSIWKMLCAL 182

```

```

Db      121 STHAYISGVSPVITILEKENDA---KYAKQNSNNELL--DDMMQHQKDIWEGMLCAL 175
QY      183 SYNTETKMKDEGVRYLYMKYIK--NNDIK--EYLEEFASRPPLRWTEMGEDFVKNR 238
      176 THKISDEERKKEIKN--KYSYKKLINESPKGSNKVDFAKKPOFLRMFTMGDEFCAQRE 232
QY      239 KELVSLKRRKCDSCITLRNNGTSNKTCDNENCGACKTQCEKYKMMERMKHHYSQKKFQ 298
      233 EKEAKVYSC-----SDAKDYDCKNTKSNASCYKACVYEDYITKKRVEYTKOKGKD 286
Db      299 LYKNSATYNNGLAVKEANSEYKNDPEVTEANSKAKHARDYLK 340
      287 AEKIT-----DKEGYEGSTNDASEYIK 309

RESULT 15
AAW00384
ID AAW00384 standard; Protein; 2913 AA.
AC AAW00384;
XX
XX 21-FEB-1997 (first entry)
XX
XX Plasmodium falciparum erythrocyte membrane protein.
XX
XX Plasmodium falciparum; erythrocyte membrane protein; malaria;
XX detection; identification; treatment; prevention; parasite.
XX
XX Plasmodium falciparum MC type.
XX
XX
XX Key Location/Qualifiers
XX Domain 62..394
XX /label= Duffy binding ligand domain 1
XX Region 607..648
XX /note="Cysteine rich motif"
XX Domain 839..1272
XX /label= Duffy binding ligand domain 2
XX Region 1482..1527
XX /note="Cysteine rich motif"
XX Domain 1706..2005
XX /label= Duffy binding ligand 3
XX Domain 2102..2349
XX /label= Duffy binding ligand 4
XX Region 2354..2398
XX /note="Cysteine rich motif"
XX Domain 2450..2475
XX /note="putative transmembrane domain"
XX
XX WO9633736-A1.
XX
XX 31-OCT-1996.
XX
XX 26-APR-1996; 96MO-US05798.
XX
XX 27-APR-1995; 95US-0430908.
XX
XX (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
XX Baruch DI, Howard RJ, Pasloske BL;
XX WPI; 1996-497376/49.
XX N-PSDB; AAT41852.
XX
XX New Plasmodium falciparum erythrocyte membrane proteins - used to
XX develop products for the diagnosis, treatment or prevention of
XX malaria parasite infections
XX
XX Claim 1; Figure 12; 149pp; English.
XX
XX A polypeptide comprising a Plasmodium falciparum (Pf) erythrocyte
XX membrane protein 1 (PFEPM1) or active fragments or analogues of that
XX protein can be used in the treatment or prevention of symptoms of a

```

CC malaria parasite infection. The polypeptides can inhibit, block or
CC reverse the sequestration of erythrocytes in patients suffering from
CC malaria. Nucleic acids derived from the p18 gene can be used as
CC probes and primers to identify a Plasmodium falciparum parasite, the
CC primers used to generate characteristic amplification patterns from
CC different P. falciparum strains. Antibodies specifically
CC immunoreactive with the p18 polypeptide or its fragments may be
CC used in diagnosis of malaria infection. This is the p18 protein
CC of the MC type of Plasmodium falciparum. An alternative, truncated
CC p18 protein is given in AAW00385.

XX
SO Sequence 2913 AA;

Query Match 21.6%; Score 412.5; DB 17; Length 2913;

Best Local Similarity 26.3%; Pred. No. 5.3e-28; Mismatches 137; Indels 165; Gaps 19;

Matches 129; Conservatively 59; Mismatches 137; Indels 165; Gaps 19;

QY 2 CKWVQKLISEOIEK-----NNIHNCKTEDEAKKWCENTKLGEEDEG--VCMPPRRONLCVH 54
DB 839 CSYIDKALKGKLDADACTLKYKGKGNNTTETSTKPGAGTSGKDTGSGICVPPRRKL--- 895
55 YLTKLND-----DSKEDLREAF 72
DB 896 YVGRILHWAGGETTEAKSOETSGGOKTPSGNESSPSSEKLPGPPETTKETPESSLHAF 955
QY 73 IKSAAAEFL---LROYNSKN-----VEDDKILHRDMIPPEFR 109
DB 956 VSPRLRLRFLLPMHFKFEQWKAQHAGAGATGOOTITIGLDGGEETPPDKLKTGHITPPDLR 1015
QY 110 SMFYTFDGYRDICL-DTDI---SEKIADHDVTA-KKITAIVFQKIGS---KTTNGK 158
DB 1016 QMFYTLGVRDILVGNNTDIYVHTSGNKKEDQIMEAIOKIEOILPTSGSSPSPPRVTOQ 1075
QY 159 KYLE--REGWKEKGLSTWKMALCALSNPTETK----MDEGRTYLMKYYKNDIKE 211
DB 1076 HSVENPRKTMWENGGKIMEGWALTYNTDTPSGTAPTOIQEVRTKIMDENSKNPKIPQ 1135
QY 212 Y-----LEEFASRPFLRWTEWGEDEFEVKNRKEIVS 243
DB 1136 KYIDQVILDDTSDAKTGTSPISGEKITPLDFTISRPPYRYLEWGEFTFCKERRKRLK 1195
QY 244 LKKKCDSCTLRNNGTSNKTG-----DDNE-----NGACKTQCEKYKMMERW 286
DB 1196 IKKEC-----RGDRGHEHCSGDGYDCTRTDADRNDFVDLNCRDCHIOCKRYKRWIDIK 1250
QY 287 KKHYSQKKKFO-----LYKNSATYNNGLAVKEANSEYKNDPEVTEANSKHAADYLYKT 341
DB 1251 FDEYHKQEKKYQGEYDKLTKDKSSGGDNCKDIEKH-----KSAAVFLK- 1295
342 QLENNMCTNG 351
DB 1296 --ELKHCKNG 1303

Search completed: June 20, 2003, 15:02:55
Job time : 59.4583 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 15:01:19 ; Search time 19.1745 Seconds
(without alignments)
538.603 Million cell updates/sec

Title: US-10-087-013-9
Perfect score: 1913
Sequence: 1 PCRMVKLISEQIEKNHN.....AKHARDYLKQLENMCTNG 351

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

arched: 262574 seqs, 29422922 residues
al number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUOS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-------------------|--------------------|
| 1 | 646 | 33.8 | 921 | 2 | US-08-568-459A-8 | Sequence 8, Appli |
| 2 | 646 | 33.8 | 921 | 2 | US-08-487-826B-8 | Sequence 8, Appli |
| 3 | 646 | 33.8 | 921 | 4 | US-09-210-288-8 | Sequence 8, Appli |
| 4 | 625.5 | 32.7 | 2710 | 2 | US-08-568-459A-12 | Sequence 12, Appli |
| 5 | 625.5 | 32.7 | 2710 | 2 | US-08-487-826B-12 | Sequence 12, Appli |
| 6 | 625.5 | 32.7 | 2710 | 4 | US-09-210-288-12 | Sequence 12, Appli |
| 7 | 625.5 | 32.7 | 3060 | 2 | US-08-487-826B-14 | Sequence 14, Appli |
| 8 | 374.5 | 19.6 | 2182 | 2 | US-08-487-826B-16 | Sequence 16, Appli |
| 9 | 353 | 18.5 | 749 | 2 | US-08-568-459A-6 | Sequence 6, Appli |
| 10 | 353 | 18.5 | 749 | 2 | US-08-487-826B-6 | Sequence 6, Appli |
| 11 | 353 | 18.5 | 749 | 4 | US-09-210-288-6 | Sequence 6, Appli |
| 12 | 329 | 17.2 | 1435 | 2 | US-08-568-459A-4 | Sequence 4, Appli |
| 13 | 329 | 17.2 | 1435 | 2 | US-08-487-826B-4 | Sequence 4, Appli |
| 14 | 329 | 17.2 | 1435 | 4 | US-09-210-288-4 | Sequence 4, Appli |
| 15 | 284 | 14.8 | 324 | 2 | US-08-568-459A-17 | Sequence 17, Appli |
| 16 | 284 | 14.8 | 324 | 2 | US-08-487-826B-29 | Sequence 29, Appli |
| 17 | 284 | 14.8 | 324 | 4 | US-09-210-288-17 | Sequence 29, Appli |
| 18 | 272.5 | 14.2 | 311 | 2 | US-08-568-459A-21 | Sequence 21, Appli |
| 19 | 272.5 | 14.2 | 311 | 2 | US-08-487-826B-33 | Sequence 33, Appli |
| 20 | 272.5 | 14.2 | 311 | 4 | US-09-210-288-21 | Sequence 21, Appli |
| 21 | 269 | 14.1 | 1115 | 2 | US-08-568-459A-2 | Sequence 2, Appli |
| 22 | 269 | 14.1 | 1115 | 2 | US-08-487-826B-2 | Sequence 2, Appli |
| 23 | 269 | 14.1 | 1115 | 4 | US-09-210-288-2 | Sequence 2, Appli |
| 24 | 231 | 12.1 | 700 | 2 | US-08-568-459A-10 | Sequence 10, Appli |
| 25 | 231 | 12.1 | 700 | 2 | US-08-487-826B-10 | Sequence 10, Appli |
| 26 | 231 | 12.1 | 700 | 2 | US-08-487-826B-10 | Sequence 10, Appli |
| 27 | 231 | 12.1 | 700 | 4 | US-09-210-288-10 | Sequence 10, Appli |

| | | | | | | |
|----|-------|-----|------|---|-------------------|--------------------|
| 28 | 171.5 | 9.0 | 778 | 6 | 5198347-4 | Patent No. 5198347 |
| 29 | 171 | 8.9 | 277 | 2 | US-08-568-459A-15 | Sequence 15, Appli |
| 30 | 171 | 8.9 | 277 | 2 | US-08-487-826B-27 | Sequence 27, Appli |
| 31 | 171 | 8.9 | 277 | 4 | US-09-210-288-15 | Sequence 15, Appli |
| 32 | 167 | 8.7 | 197 | 6 | 5198347-2 | Patent No. 5198347 |
| 33 | 144.5 | 7.6 | 411 | 2 | US-08-568-459A-20 | Sequence 20, Appli |
| 34 | 144.5 | 7.6 | 411 | 2 | US-08-487-826B-32 | Sequence 32, Appli |
| 35 | 144.5 | 7.6 | 411 | 4 | US-09-210-288-20 | Sequence 20, Appli |
| 36 | 137 | 7.2 | 282 | 2 | US-08-568-459A-16 | Sequence 16, Appli |
| 37 | 137 | 7.2 | 282 | 2 | US-08-487-826B-28 | Sequence 28, Appli |
| 38 | 137 | 7.2 | 282 | 4 | US-09-210-288-16 | Sequence 16, Appli |
| 39 | 135 | 7.1 | 291 | 2 | US-08-568-459A-13 | Sequence 13, Appli |
| 40 | 135 | 7.1 | 291 | 2 | US-08-487-826B-25 | Sequence 25, Appli |
| 41 | 135 | 7.1 | 291 | 4 | US-09-210-288-13 | Sequence 13, Appli |
| 42 | 131.5 | 6.9 | 362 | 2 | US-08-568-459A-18 | Sequence 18, Appli |
| 43 | 131.5 | 6.9 | 362 | 2 | US-08-487-826B-30 | Sequence 30, Appli |
| 44 | 131.5 | 6.9 | 362 | 4 | US-09-210-288-18 | Sequence 18, Appli |
| 45 | 107 | 5.6 | 1507 | 3 | US-08-929-329-5 | Sequence 5, Appli |

ALIGNMENTS

RESULT 1
US-08-568-459A-8
: Sequence 8, Application US/08568459A
: Patent No. 5849306
: GENERAL INFORMATION:
: APPLICANT: Sim, Kim L.
: APPLICANT: Chitnis, Chetan
: APPLICANT: Miller, Louis H.
: APPLICANT: Peterson, David S.
: APPLICANT: Su, Xin-zhan
: APPLICANT: Wellens, Thomas E.
: TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
: TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Knobbe Martens Olson & Bear
: STREET: 620 Newport Center Drive 16th Floor
: CITY: Newport Beach
: STATE: California
: COUNTRY: US
: ZIP: 92660
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/568,459A
: FILING DATE: 07-DEC-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Israel, Ned
: REGISTRATION NUMBER: 29,655
: REFERENCE/DOCKET NUMBER: NIH121.001CPI
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 235-8550
: TELEFAX: (619) 235-0176
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 921 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Plasmodium falciparum
: US-08-568-459A-8
Query Match 33.8%, Score 646, DB 2, Length 921;

Best Local Similarity 41.0%; Pred. No. 2.4e-50;
Matches 150; Conservative 54; Mismatches 112; Indels 50; Gaps 16;

```
QY 2 CKMVKQLISEQLEKNNIHNCKTDE-AKMKCENT--KLGDEGVCMPPRRONLCVHYLTK 58
Db 396 CEIVAMLMKDKNGRTTVEGCRKEYSEMTCDSESKIMQO-HGACIPPRKOLCLHYLER 454
QY 59 LNDKSEEDLRPAFISAAAEFLRLROY--NSKNVED-DKILHRDIPPEFFRSMEYT 114
Db 455 IMTNTNE--LKYAFICAAAEFLRLQNYKKDKNGMAEDLEKLGKIIPEDPKROMFT 512
QY 115 FGDRDIDLDPDI-SEKIDHDVYTTAKKKTITAVFOKISGKTTNGKKVLEREGMKKEYGLS 173
Db 513 FADYRDLGTDISSKKDTISKGVGKVCNIDVFKISN-----SIRYRKSMWETNGPV 566
QY 174 IMKGMICALSYNT-----ET-KKMDGVRTYLMKYKNNDIKEYLEEFASPPFLRW 225
Db 567 IMEGMICALSYDTLSLNNVPETHKKLTGECNNNF-EKVIF-GSDSSTLSKSEKPOFLRW 624
QY 226 VTEMGEDFVKNRKKEVLSLKKKCDSCITLRNNGTSNKTCDNENCGACKTQCEKRYKKMER 285
Db 625 LTEWGENFCKEOKREKRYLLAKCKDQVDGDC---KC--NGKCVACKDCKOYHSMIGI 678
QY 286 WKHYSSQKKFQLYKNSATYNNGLAVKANSEYTKNDPEVTEANSKARADYLTQLEN 345
Db 679 WIDNKKKQKGRYTEVK-----KIPLYEKEDVKNSDD--ARDYLTQLEN 721
QY 346 MICTNG 351
Db 722 MKCVNG 727
```

RESULT 2

US-08-487-826B-8
Sequence 8, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Istraelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 921 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-08-487-826B-8

Query Match 33.8%; Score 646; DB 2; Length 921;
Best Local Similarity 41.0%; Pred. No. 2.4e-50;
Matches 150; Conservative 54; Mismatches 112; Indels 50; Gaps 16;

```
QY 2 CKMVKQLISEQLEKNNIHNCKTDE-AKMKCENT--KLGDEGVCMPPRRONLCVHYLTK 58
Db 396 CEIVAMLMKDKNGRTTVEGCRKEYSEMTCDSESKIMQO-HGACIPPRKOLCLHYLER 454
QY 59 LNDKSEEDLRPAFISAAAEFLRLROY--NSKNVED-DKILHRDIPPEFFRSMEYT 114
Db 455 IMTNTNE--LKYAFICAAAEFLRLQNYKKDKNGMAEDLEKLGKIIPEDPKROMFT 512
QY 115 FGDRDIDLDPDI-SEKIDHDVYTTAKKKTITAVFOKISGKTTNGKKVLEREGMKKEYGLS 173
Db 513 FADYRDLGTDISSKKDTISKGVGKVCNIDVFKISN-----SIRYRKSMWETNGPV 566
QY 174 IMKGMICALSYNT-----ET-KKMDGVRTYLMKYKNNDIKEYLEEFASPPFLRW 225
Db 567 IMEGMICALSYDTLSLNNVPETHKKLTGECNNNF-EKVIF-GSDSSTLSKSEKPOFLRW 624
QY 226 VTEMGEDFVKNRKKEVLSLKKKCDSCITLRNNGTSNKTCDNENCGACKTQCEKRYKKMER 285
Db 625 LTEWGENFCKEOKREKRYLLAKCKDQVDGDC---KC--NGKCVACKDCKOYHSMIGI 678
QY 286 WKHYSSQKKFQLYKNSATYNNGLAVKANSEYTKNDPEVTEANSKARADYLTQLEN 345
Db 679 WIDNKKKQKGRYTEVK-----KIPLYEKEDVKNSDD--ARDYLTQLEN 721
QY 346 MICTNG 351
Db 722 MKCVNG 727
```

RESULT 3

US-09-210-288-8
Sequence 8, Application US/09210288
Patent No. 6392026

GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

NAME: Fuller, Michael

MEDIUM TYPE: Floppy disk

STREET: 620 NEWPORT CENTER DRIVE 10TH FLOOR
CITY: Newport Beach

```

STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-08-487-826B-12

```

```

Query Match          32.7%; Score 625.5; DB 2; Length 2710;
Best Local Similarity 38.8%; Pred. No. 8.1e-48;
Matches 137; Conservative 61; Mismatches 126; Indels 29; Gaps 9;

QY 2 CKWVOKLISEQIEKNNIHNCKTDED-AKMKCENTLGEDEGCMPPRROMLCVHYL---T 57
DB 1349 CKIVEKILEGKNGRTTVEGCNPKESYPDMDCNNIDISHGACMPBRROKLCLYIAHES 1408
QY 58 KLNDSKEEDLEAFIKSAAEFTLLRQYNSKNVEDDKILHMDIPEEFFRSMFTYFGD 117
DB 1409 QTENIKITDNLKDAFIKTAAEFTLSWQYKSKNDSSEAKILLDGLIPSOFLRSMYTFGD 1468
QY 118 YRDICLDTISEKIDHDTTAKKKTITAVFQKIGSKTTNGKVLREGEWKEVLSIWKG 177
DB 1469 YRDICLNTDISK--QNDVAKAKDKIGKFEKSGSKSPSG---LSRQEMWKTNGPELWKG 1523
QY 178 MLCALS-YNTEETKMDGCVTYLMKYIYKNDIKEYLEEFASRPPEFLRWVTEWGEDFVKN 236
DB 1524 MLCALTKRYVTDITNKRRIKNDYSYDKVNOGNGNPISLEEFAPKQFLRMHTEGEEFCAE 1583
QY 237 RKKEIVSLKKKCDSCITLRNNGTSNKTCDNENGCACKTQCEKRYKMERMKKHYSQKK 296
DB 1584 ROKKENITIDAC-----NEINSTQOCNDAKH--RCNOACRAVOEYEVENKKKEFSGQTNN 1635
QY 297 FOLYKNSATYNNGLAVKEANSETYKNDPEVTEANSAKHARDYLKTOLENMICT 349
DB 1636 FVLKANVQPOD-----PEYKGYEYKGVQPIQGN-----EYLLQKCDNNKCS 1677

```

```

RESULT 6
US-09-210-288-12
Sequence 12, Application US/09210288
Patent No. 6392026
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

```

```

NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210, 288
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-09-210-288-12

```

```

Query Match          32.7%; Score 625.5; DB 4; Length 2710;
Best Local Similarity 38.8%; Pred. No. 8.1e-48;
Matches 137; Conservative 61; Mismatches 126; Indels 29; Gaps 9;

QY 2 CKWVOKLISEQIEKNNIHNCKTDED-AKMKCENTLGEDEGCMPPRROMLCVHYL---T 57
DB 1349 CKIVEKILEGKNGRTTVEGCNPKESYPDMDCNNIDISHGACMPBRROKLCLYIAHES 1408
QY 58 KLNDSKEEDLEAFIKSAAEFTLLRQYNSKNVEDDKILHMDIPEEFFRSMFTYFGD 117
DB 1409 QTENIKITDNLKDAFIKTAAEFTLSWQYKSKNDSSEAKILLDGLIPSOFLRSMYTFGD 1468
QY 118 YRDICLDTISEKIDHDTTAKKKTITAVFQKIGSKTTNGKVLREGEWKEVLSIWKG 177
DB 1469 YRDICLNTDISK--QNDVAKAKDKIGKFEKSGSKSPSG---LSRQEMWKTNGPELWKG 1523
QY 178 MLCALS-YNTEETKMDGCVTYLMKYIYKNDIKEYLEEFASRPPEFLRWVTEWGEDFVKN 236
DB 1524 MLCALTKRYVTDITNKRRIKNDYSYDKVNOGNGNPISLEEFAPKQFLRMHTEGEEFCAE 1583
QY 237 RKKEIVSLKKKCDSCITLRNNGTSNKTCDNENGCACKTQCEKRYKMERMKKHYSQKK 296
DB 1584 ROKKENITIDAC-----NEINSTQOCNDAKH--RCNOACRAVOEYEVENKKKEFSGQTNN 1635
QY 297 FOLYKNSATYNNGLAVKEANSETYKNDPEVTEANSAKHARDYLKTOLENMICT 349
DB 1636 FVLKANVQPOD-----PEYKGYEYKGVQPIQGN-----EYLLQKCDNNKCS 1677

```

```

RESULT 7
US-08-487-826B-14
Sequence 14, Application US/08487826B
Patent No. 593827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.

```

```

?
? APPLICANT: Peterson, David S.
? APPLICANT: Su, Xin-zhaun
? APPLICANT: Wellens, Thomas E.
? TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
? NUMBER OF SEQUENCES: 45
? CORRESPONDENCE ADDRESS:
? ADDRESS: Knobbe Martens Olson & Bear
? STREET: 620 Newport Center Drive 16th Floor
? CITY: Newport Beach
? STATE: California
? COUNTRY: US
? ZIP: 92660
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/487,826B
? FILING DATE: 10-SEP-1993
? CLASSIFICATION: 435
?
? ATTORNEY/AGENT INFORMATION:
? NAME: Israelson, Ned
? REGISTRATION NUMBER: 29,655
? REFERENCE/DOCKET NUMBER: NIH121.001CPI
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619) 235-8550
? TELEFAX: (619) 235-0176
? INFORMATION FOR SEQ ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 3060 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
?
? US-08-487-826B-14
?
?
? Query Match 32.7%; Score 625.5; DB 2; Length 3060;
? Best Local Similarity 38.8%; Pred. No. 9.7e-48;
? Matches 137; Conservative 61; Mismatches 126; Indels 29; Gaps 9;
?
? QY 2 CKNVOKLISEQIEKNNIHNCKTDED-AKMKCENTKLGDEGVCMPPRRQMLCVHYL--T 57
? DB 1347 CKIVKILBCKNRTYVGECPNPKESYPMDCNNIDISHGACMPRRQMLCVHYAHES 1406
?
? 58 KLANDSKEDLDREAFIKSAAEFPLRLQYNSKNVEDDKTLHRDMIPPEFFRSMEYTFEGD 117
? DB 1407 QTENIKRTDDMLKDAFIKTAAEFTLSWQYKSKNDSEAKILDRGLIPSOFLRSMTYFGD 1466
?
? 118 YRDICLDTDISEKTIADHDVTTAKKITAFAVQKIGSKTNGKTVLERGMMKREYGLSTWKG 177
? DB 1467 YRICLINTDLSKR--QNDVAKADKIGKFSKDSKSPSG---LSROEMKKTGPELWKG 1521
?
? QY 178 MLCALS-YNTETKKMDEGVFTYLMKYYIKNNDIKEYLEEFASRPPLRWVTEWGEDFVK 236
? DB 1522 MLCALIRKYVDTDNKRIRKNDYSYDKVNGSONGNSPLEEFAKPOFLRWMIEMGEBCAE 1581
?
? QY 237 RKKEIYSLKKKCSCLRNNGTSNKTCDNDNENGACACTQCEKTKKMMERKKHHYSKKR 236
? DB 1582 ROKKENTIKDAC-----NEINSTQOCNDAKH--RCNOACRAVOYEENKKEESGQTN 1633
?
? QY 297 FOYKKSATYNNGLAVEANSSEYKNDPEYTEANSAKHARDYLTOLENNICT 349
? DB 1634 FVLKANVQPOD-----PEYKGYEKDGVQPIQGN-----EYLLQKCDNNKCS 1675
?
?
? RESULT 8
? US-08-487-826B-16
? Sequence 16, Application US/08487826B
? Patent No. 5993827
? GENERAL INFORMATION:
? APPLICANT: Sim, Kim L.

```

```

?
? APPLICANT: Chitnis, Chetan
? APPLICANT: Miller, Louis H.
? APPLICANT: Peterson, David S.
? APPLICANT: Su, Xin-zhaun
? APPLICANT: Wellens, Thomas E.
? TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
? NUMBER OF SEQUENCES: 45
? CORRESPONDENCE ADDRESS:
? ADDRESS: Knobbe Martens Olson & Bear
? STREET: 620 Newport Center Drive 16th Floor
? CITY: Newport Beach
? STATE: California
? COUNTRY: US
? ZIP: 92660
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/487,826B
? FILING DATE: 10-SEP-1993
? CLASSIFICATION: 435
?
? ATTORNEY/AGENT INFORMATION:
? NAME: Israelson, Ned
? REGISTRATION NUMBER: 29,655
? REFERENCE/DOCKET NUMBER: NIH121.001CPI
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619) 235-8550
? TELEFAX: (619) 235-0176
? INFORMATION FOR SEQ ID NO: 16:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2182 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? FRAGMENT TYPE: N-terminal
?
? US-08-487-826B-16
?
?
? Query Match 19.6%; Score 374.5; DB 2; Length 2182;
? Best Local Similarity 25.4%; Pred. No. 4.4e-25;
? Matches 121; Conservative 54; Mismatches 137; Indels 165; Gaps 19;
?
? QY 1 PCKMVOKLISEQIEKNNIHNCK-----KKTEDAKMKC-----ENTKLGDEGVCMPPR 47
? DB 851 PCQIVGKLFED--DKSLKEACGLKYGGRKEKFPNMKCVTPSGVSTATSGDGLCVPR 908
?
? 48 RQNLVCVHYL-----TKLNDKSKEDLDREAFIKSAAEFTPLRLQYNS 89
? DB 909 RRLRYVGLSQWASRGDETEVSSSEATSPASQSESEKLTALTESAITEFTPLMHRYKE 968
?
? QY 90 KNV-----ED--DKILHRDMIPPEFFRSMEYTFEGDYRDICL--D 124
? DB 969 EKKRPATQDAGLGVSLPSPSPGEDQQTQLOQTGVLPPLFLRMFTTLADYDLXSGS 1028
?
? QY 125 TDISEKTIADHDVTTAK-----KKTAVFQKIGSKYTN 156
? DB 1029 NDTSD-----TTGKQTPSSSNDNLKNIIVLEASGSTOEKEMKKQIQAKIKTLNGATS 1081
?
? QY 157 GKRYLER-----EGWKKREYGLSIWGMICALSY-----NTERKK--MDE 193
? DB 1082 GVPVYTKNSVYKTPQQTWMMENIANDIWMWMCALTYKENDARGTSAKTEQKKDKLKAALME 1141
?
? QY 194 GVR-TYLMKYYI-----KNNDIKE--YLEEFASRPPLRWVTEWGEDFVKNRKK 239
? DB 1142 ANKNTPLEKYQYTNVNLKLEDSGAKSNDTIQPTPLKNEVELPTFRWLHBMGNSGFFRRK 1201
?
? QY 240 ELVSLKKKCSCLLRNNGTSNKTCDNENC-----GACKTQCEKYYKKWME 284

```

Db 1202 RLAQIKHEC-----MDEDEKOYSGDEYCEEIFSKOYNVLQDLSSSCAKPCLYKWTIE 1256
QY 285 RPKKHYSQKKKFKLYKNSATYNNGLAVKANSEIYKNDPEYTEANSKADHDYLT 341
Db 1257 KKKTEKQKQKAVEQK-----SNTEHQKQKQKOTQSN--NAMEFSHT 1298

RESULT 9
US-08-568-459A-6
; Sequence 6, Application US/08568459A
; Patent No. 5849306
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelien, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 749 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; US-08-568-459A-6

Query Match 18.5%; Score 353; DB 2; Length 749;
Best Local Similarity 28.2%; Pred. No. 8.9e-24;
Matches 103; Conservative 50; Mismatches 118; Indels 94; Gaps 16;

QY 10 SEQIEKNNHNCK---KTDAAKMC-ENTKLGDEGVCMPRRQNCVHLYTLKNDLSKE 65
Db 103 SKEHESSVFGCKTKISKVKKMKWCYSNNKVTKEGVCGRPRQOLGVIPLIR-DGNE 161
QY 66 EDLEAFIKSAAETFLRLQYNSKNVEDKILHRMIPPEFRSMFTYRGDRDLCIDT 125
Db 162 EGLKDHINKANAEAMLEKXY--ENAGGDKICN-----AIIGSTADIGDIVRGL 209
QY 126 DIEKADHDVYTAKKRTAVFQKISGKTNGKRVL---EREGWMEYGLSTWKMGLCA 181
Db 210 DVM-----RDINT--NKLSEKFORIFMGSGNSRRKKQNDNERNKKNWKKQNLWSSM--- 259
QY 182 LSVNTEFKMDEGVRYLWKYIIRKNDIKELYLEEFASRPFLWVTEWGEDVFNKRRKL 241

Db 260 -----VKHLPKGTCKRH--NNEKIPDPLRWLKEWGDEFCCEMGTEY 300
QY 242 VSLKRCQSCITLRNNGTSNKTCDDNENGACACKTQCEKFKKMERKRIHYSQKKF----- 297
Db 301 KOLEKICE-----NKNCKSEK-----CKNACSSYEKIKERKNYINQSKKFDSDK 346
QY 298 QLYKNSATYNNGLAVKANSEIYK-----NDPEYTEANSKADHDYLTQLEN 345
Db 347 KLKKRNNLYN-----KFEDESKAYLRSESKQCSNIEFNDEFTFPKRYKEA-----C 392
QY 346 MICTN 350
Db 393 WVCEN 397

RESULT 10
US-08-487-826B-6
; Sequence 6, Application US/08487826B
; Patent No. 5938827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelien, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 749 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; US-08-487-826B-6

Query Match 18.5%; Score 353; DB 2; Length 749;
Best Local Similarity 28.2%; Pred. No. 8.9e-24;
Matches 103; Conservative 50; Mismatches 118; Indels 94; Gaps 16;

QY 10 SEQIEKNNHNCK---KTDAAKMC-ENTKLGDEGVCMPRRQNCVHLYTLKNDLSKE 65
Db 103 SKEHESSVFGCKTKISKVKKMKWCYSNNKVTKEGVCGRPRQOLGVIPLIR-DGNE 161


```

01 Yy 66 EDIREARIKAAEFLLRLROYNSKVNEDDKILHRMIDPEFFRSNFYFGDYRDCIDT 125
02 Yy 162 EGLKDHNRKANAEAMHLKERY--EAGGDKICN-----ALIGSYADIGIVAGL 209
03 Yy 126 -DISIKADIADHVTTAKKRITAVFOKISGTINGRKVL---EREGRWKEYGILSIWGMJCA 181
04 Yy 210 DVM-----RDINT--NKLSKEFKQIFMGGSNSRKKNDNNERNKRWKQORNLIWESM--- 259
05 Yy 182 LSVNTEFKMKHDEGVRYLKKIYIKKNDIYELEEFPSRPPLRVWTEWGEDVYVKNRKKEL 241
06 Yy 260 -----VKHIDPKYTKCRH--NNEFKIPOFLRWLKEWDEFCCEEMGTVEV 300
07 Yy 242 VSLKKKDCSTLLRNNGTSNKTCDNNGCAGKTCQCEKYKKWMBRWKHHYSOKKRF---- 297
08 Yy 301 KOLEKICE-----KNCSEKK-----CKNACGSYKWKIKERKNEVNLQSKKFDSDX 346
09 Yy 298 QLYKNSATYNGLAIVKAEANSEYTK-----NDEVEYFANSAKARADYLTQLEN 345
10 Yy 347 KLNKKNNLYN----KFEDSKAYLRSESKQCSNIEFNDETFTFPNKKYEA-----C 392
11 Yy 346 MICTN 350
12 Yy 393 MWCEN 397
13 Yy
14 Yy
15 Yy
16 Yy
17 Yy
18 Yy
19 Yy
20 Yy
21 Yy
22 Yy
23 Yy
24 Yy
25 Yy
26 Yy
27 Yy
28 Yy
29 Yy
30 Yy
31 Yy
32 Yy
33 Yy
34 Yy
35 Yy
36 Yy
37 Yy
38 Yy
39 Yy
40 Yy
41 Yy
42 Yy
43 Yy
44 Yy
45 Yy
46 Yy
47 Yy
48 Yy
49 Yy
50 Yy
51 Yy
52 Yy
53 Yy
54 Yy
55 Yy
56 Yy
57 Yy
58 Yy
59 Yy
60 Yy
61 Yy
62 Yy
63 Yy
64 Yy
65 Yy
66 Yy
67 Yy
68 Yy
69 Yy
70 Yy
71 Yy
72 Yy
73 Yy
74 Yy
75 Yy
76 Yy
77 Yy
78 Yy
79 Yy
80 Yy
81 Yy
82 Yy
83 Yy
84 Yy
85 Yy
86 Yy
87 Yy
88 Yy
89 Yy
90 Yy
91 Yy
92 Yy
93 Yy
94 Yy
95 Yy
96 Yy
97 Yy
98 Yy
99 Yy
100 Yy
101 Yy
102 Yy
103 Yy
104 Yy
105 Yy
106 Yy
107 Yy
108 Yy
109 Yy
110 Yy
111 Yy
112 Yy
113 Yy
114 Yy
115 Yy
116 Yy
117 Yy
118 Yy
119 Yy
120 Yy
121 Yy
122 Yy
123 Yy
124 Yy
125 Yy
126 Yy
127 Yy
128 Yy
129 Yy
130 Yy
131 Yy
132 Yy
133 Yy
134 Yy
135 Yy
136 Yy
137 Yy
138 Yy
139 Yy
140 Yy
141 Yy
142 Yy
143 Yy
144 Yy
145 Yy
146 Yy
147 Yy
148 Yy
149 Yy
150 Yy
151 Yy
152 Yy
153 Yy
154 Yy
155 Yy
156 Yy
157 Yy
158 Yy
159 Yy
160 Yy
161 Yy
162 Yy
163 Yy
164 Yy
165 Yy
166 Yy
167 Yy
168 Yy
169 Yy
170 Yy
171 Yy
172 Yy
173 Yy
174 Yy
175 Yy
176 Yy
177 Yy
178 Yy
179 Yy
180 Yy
181 Yy
182 Yy
183 Yy
184 Yy
185 Yy
186 Yy
187 Yy
188 Yy
189 Yy
190 Yy
191 Yy
192 Yy
193 Yy
194 Yy
195 Yy
196 Yy
197 Yy
198 Yy
199 Yy
200 Yy
201 Yy
202 Yy
203 Yy
204 Yy
205 Yy
206 Yy
207 Yy
208 Yy
209 Yy
210 Yy
211 Yy
212 Yy
213 Yy
214 Yy
215 Yy
216 Yy
217 Yy
218 Yy
219 Yy
220 Yy
221 Yy
222 Yy
223 Yy
224 Yy
225 Yy
226 Yy
227 Yy
228 Yy
229 Yy
230 Yy
231 Yy
232 Yy
233 Yy
234 Yy
235 Yy
236 Yy
237 Yy
238 Yy
239 Yy
240 Yy
241 Yy
242 Yy
243 Yy
244 Yy
245 Yy
246 Yy
247 Yy
248 Yy
249 Yy
250 Yy
251 Yy
252 Yy
253 Yy
254 Yy
255 Yy
256 Yy
257 Yy
258 Yy
259 Yy
260 Yy
261 Yy
262 Yy
263 Yy
264 Yy
265 Yy
266 Yy
267 Yy
268 Yy
269 Yy
270 Yy
271 Yy
272 Yy
273 Yy
274 Yy
275 Yy
276 Yy
277 Yy
278 Yy
279 Yy
280 Yy
281 Yy
282 Yy
283 Yy
284 Yy
285 Yy
286 Yy
287 Yy
288 Yy
289 Yy
290 Yy
291 Yy
292 Yy
293 Yy
294 Yy
295 Yy
296 Yy
297 Yy
298 Yy
299 Yy
300 Yy
301 Yy
302 Yy
303 Yy
304 Yy
305 Yy
306 Yy
307 Yy
308 Yy
309 Yy
310 Yy
311 Yy
312 Yy
313 Yy
314 Yy
315 Yy
316 Yy
317 Yy
318 Yy
319 Yy
320 Yy
321 Yy
322 Yy
323 Yy
324 Yy
325 Yy
326 Yy
327 Yy
328 Yy
329 Yy
330 Yy
331 Yy
332 Yy
333 Yy
334 Yy
335 Yy
336 Yy
337 Yy
338 Yy
339 Yy
340 Yy
341 Yy
342 Yy
343 Yy
344 Yy
345 Yy
346 Yy
347 Yy
348 Yy
349 Yy
350 Yy
351 Yy
352 Yy
353 Yy
354 Yy
355 Yy
356 Yy
357 Yy
358 Yy
359 Yy
360 Yy
361 Yy
362 Yy
363 Yy
364 Yy
365 Yy
366 Yy
367 Yy
368 Yy
369 Yy
370 Yy
371 Yy
372 Yy
373 Yy
374 Yy
375 Yy
376 Yy
377 Yy
378 Yy
379 Yy
380 Yy
381 Yy
382 Yy
383 Yy
384 Yy
385 Yy
386 Yy
387 Yy
388 Yy
389 Yy
390 Yy
391 Yy
392 Yy
393 Yy
394 Yy
395 Yy
396 Yy
397 Yy
398 Yy
399 Yy
400 Yy
401 Yy
402 Yy
403 Yy
404 Yy
405 Yy
406 Yy
407 Yy
408 Yy
409 Yy
410 Yy
411 Yy
412 Yy
413 Yy
414 Yy
415 Yy
416 Yy
417 Yy
418 Yy
419 Yy
420 Yy
421 Yy
422 Yy
423 Yy
424 Yy
425 Yy
426 Yy
427 Yy
428 Yy
429 Yy
430 Yy
431 Yy
432 Yy
433 Yy
434 Yy
435 Yy
436 Yy
437 Yy
438 Yy
439 Yy
440 Yy
441 Yy
442 Yy
443 Yy
444 Yy
445 Yy
446 Yy
447 Yy
448 Yy
449 Yy
450 Yy
451 Yy
452 Yy
453 Yy
454 Yy
455 Yy
456 Yy
457 Yy
458 Yy
459 Yy
460 Yy
461 Yy
462 Yy
463 Yy
464 Yy
465 Yy
466 Yy
467 Yy
468 Yy
469 Yy
470 Yy
471 Yy
472 Yy
473 Yy
474 Yy
475 Yy
476 Yy
477 Yy
478 Yy
479 Yy
480 Yy
481 Yy
482 Yy
483 Yy
484 Yy
485 Yy
486 Yy
487 Yy
488 Yy
489 Yy
490 Yy
491 Yy
492 Yy
493 Yy
494 Yy
495 Yy
496 Yy
497 Yy
498 Yy
499 Yy
500 Yy
501 Yy
502 Yy
503 Yy
504 Yy
505 Yy
506 Yy
507 Yy
508 Yy
509 Yy
510 Yy
511 Yy
512 Yy
513 Yy
514 Yy
515 Yy
516 Yy
517 Yy
518 Yy
519 Yy
520 Yy
521 Yy
522 Yy
523 Yy
524 Yy
525 Yy
526 Yy
527 Yy
528 Yy
529 Yy
530 Yy
531 Yy
532 Yy
533 Yy
534 Yy
535 Yy
536 Yy
537 Yy
538 Yy
539 Yy
540 Yy
541 Yy
542 Yy
543 Yy
544 Yy
545 Yy
546 Yy
547 Yy
548 Yy
549 Yy
550 Yy
551 Yy
552 Yy
553 Yy
554 Yy
555 Yy
556 Yy
557 Yy
558 Yy
559 Yy
560 Yy
561 Yy
562 Yy
563 Yy
564 Yy
565 Yy
566 Yy
567 Yy
568 Yy
569 Yy
570 Yy
571 Yy
572 Yy
573 Yy
574 Yy
575 Yy
576 Yy
577 Yy
578 Yy
579 Yy
580 Yy
581 Yy
582 Yy
583 Yy
584 Yy
585 Yy
586 Yy
587 Yy
588 Yy
589 Yy
590 Yy
591 Yy
592 Yy
593 Yy
594 Yy
595 Yy
596 Yy
597 Yy
598 Yy
599 Yy
600 Yy
601 Yy
602 Yy
603 Yy
604 Yy
605 Yy
606 Yy
607 Yy
608 Yy
609 Yy
610 Yy
611 Yy
612 Yy
613 Yy
614 Yy
615 Yy
616 Yy
617 Yy
618 Yy
619 Yy
620 Yy
621 Yy
622 Yy
623 Yy
624 Yy
625 Yy
626 Yy
627 Yy
628 Yy
629 Yy
630 Yy
631 Yy
632 Yy
```

```

Query March 18.5%: Score 353; DB 4; Length 749;
Best Local Similarity 28.2%: Pred No. 8.9e-24;
Matches 103; Conservative 50; Mismatches 118; Indels 94; Gaps 16;

QY 10 SEQLEKNNINCK---KTEDAKWK-C-ENFKLGEDGVCMPPRNMLCVHYLTKLNDSDKE 65
   | : : : : | | | : : : : | | | : : : : |
Db 103 SKHEESSVGCOTKISKYKKKWNVCYSNNKVKTPKEGVCGRPQQLGGLYFILR-DGNE 161
   | : : : : | | | : : : : | | | : : : : |
QY 66 EDLREAFIKSAAEFTLLRQYYSKKNVEDDKILHRLMTPPEFRFSMYTFGGYRQICDT 125
   | : : : : | | | : : : : | | | : : : : |
Db 162 EGGLEKDHINKAANYEAMHLKERY--ENAGGDKICN-----ALIGYADIGILVRL 209
   | : : : : | | | : : : : | | | : : : : |
QY 126 DISKELADHVVTAKKKIKITAVFQKISKTNGCKVY-----EREGWKKREGLSTWGMGLCA 181
   | : : : : | | | : : : : | | | : : : : |
Db 210 DVM-----RQINT--NKLSEFKPIKIMWGSGNSKKNDNNERNKKMKWKKERNLWSSM--- 259
   | : : : : | | | : : : : | | | : : : : |
QY 182 LSYNTEFKKMDGCVRYLMKRYIKKNDIKYEELFEFASRPPLRWWTWEGEDFVKNKKEL 241
   | : : : : | | | : : : : | | | : : : : |
Db 260 -----VKHIFPKGTYCKRH--NNEFKIIPQFLRWLKEKDEFCCEEGTEV 300
   | : : : : | | | : : : : | | | : : : : |
QY 242 VSLKKKDCSLTRNNGTSNKTCDQDNENCGCKTQCKRYKKMMERMKKHHSQKKF----- 297
   | : : : : | | | : : : : | | | : : : : |
Db 301 KOLEKIE-----NKNQSEKK---CKNACSSYERKWIKERENENLQSKKFDSDK 346
   | : : : : | | | : : : : | | | : : : : |
QY 298 OLYKNSATYNNGLAVKFEANSEYK-----NDEVTVEANSKARHARDYLYKTQLEBN 345
   | : : : : | | | : : : : | | | : : : : |
Db 347 KLNKKNNLYN-----KEDSKATLRSESKOCSNIEFNDELFTTPPKYKEL-----C 392
   | : : : : | | | : : : : | | | : : : : |
QY 346 MICTN 350
   | : : : : |
Db 393 MVCEN 397

RESULT 12
US-08-568-459A-4
Sequence 4, Application US/08568459A
Patent No. 5849306
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chltnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568.459A
FILING DATE: 07-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

```

LENGTH: 1435 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-08-568-459A-4

Query Match 17.2%; Score 329; DB 2; Length 1435;
Best Local Similarity 24.8%; Pred No. 3.4e-21;
Matches 87; Conservative 63; Mismatches 113; Indels 88; Gaps 13;

QY 14 EKNNIHNCRTEDAKKCENTKLGDEGVCMPPRRONLCVHYLTKLNDSS---KEEDLR 69
DB 463 DKNSV---DTNKTWCKKNPYLSTKDYCVPRRQELCLGNIDRIYKNLMIKHHILA 518
QY 70 EAFIKSAAEFTLLRQYNSKKNVEDKILHRDMMIPPEFRSMFTYFGYRIDCLDTDISE 129
DB 519 IAIYES---RLIKRRYKKNKDKDEVCIIKK-----TFADIRDIIGGTGYWN 561
DB 130 KIADHDVTTAKKITAVFOKIGSKTTNGKKV-----LEREGWKEEGLSIWGMCA 181
DB 562 DLSNRKL-----VGKINTNSKYVHRNKNKDKLPRDEWKKYIKKDYWN----- 603
QY 182 LSYNETKMDGCVRTYLMKTIYKNNDI-KEYLEEFASRPPPLRWVTGEGDFVKNRKE 240
DB 604 -----VISWFKDITYCKE--DDIENIPQEFFRWFSEMGDDYCODKTKM 644
QY 241 LVSLKKDCSCTLRNNGTSNKTCDNENCGACKTCCEYKKMKMEKKHYSOKKFFOLY 300
DB 645 IETLKEC-----KEKCEDDN-----CKSKCNSTKEMISKKEEYKOKAYOYEY 690
QY 301 KNSATYNNGLAVKEANSEY--KNDPEVTEANSKHAHARDYLTQLENNICT 349
DB 691 QKGNMYKMYSEKSIKPEVYLLKYSKCSNLNFEDEFKEELHSDYKNK-CT 740

RESULT 13

US-08-487-826B-4
Sequence 4, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487, 826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI

TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1435 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-08-487-826B-4

Query Match 17.2%; Score 329; DB 2; Length 1435;
Best Local Similarity 24.8%; Pred No. 3.4e-21;
Matches 87; Conservative 63; Mismatches 113; Indels 88; Gaps 13;

QY 14 EKNNIHNCRTEDAKKCENTKLGDEGVCMPPRRONLCVHYLTKLNDSS---KEEDLR 69
DB 463 DKNSV---DTNKTWCKKNPYLSTKDYCVPRRQELCLGNIDRIYKNLMIKHHILA 518
QY 70 EAFIKSAAEFTLLRQYNSKKNVEDKILHRDMMIPPEFRSMFTYFGYRIDCLDTDISE 129
DB 519 IAIYES---RLIKRRYKKNKDKDEVCIIKK-----TFADIRDIIGGTGYWN 561
DB 130 KIADHDVTTAKKITAVFOKIGSKTTNGKKV-----LEREGWKEEGLSIWGMCA 181
DB 562 DLSNRKL-----VGKINTNSKYVHRNKNKDKLPRDEWKKYIKKDYWN----- 603
QY 182 LSYNETKMDGCVRTYLMKTIYKNNDI-KEYLEEFASRPPPLRWVTGEGDFVKNRKE 240
DB 604 -----VISWFKDITYCKE--DDIENIPQEFFRWFSEMGDDYCODKTKM 644
QY 241 LVSLKKDCSCTLRNNGTSNKTCDNENCGACKTCCEYKKMKMEKKHYSOKKFFOLY 300
DB 645 IETLKEC-----KEKCEDDN-----CKSKCNSTKEMISKKEEYKOKAYOYEY 690
QY 301 KNSATYNNGLAVKEANSEY--KNDPEVTEANSKHAHARDYLTQLENNICT 349
DB 691 QKGNMYKMYSEKSIKPEVYLLKYSKCSNLNFEDEFKEELHSDYKNK-CT 740

RESULT 14

US-09-210-286-4
Sequence 4, Application US/09210286
Patent No. 6392026

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210, 286
FILING DATE:

```
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1435 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum

Query Match
Best local Similarity 17.2%; Score 329; DB 4; Length 1435;
Matches 87; Conservative 63; Mismatches 113; Indels 88; Gaps 13;

QY 14 EKNNHNCCKTEDAKKCENTKLGEDGVCMPPRONLCVHYLTKLNDSD---KEEDLR 69
DB 463 DKNSV---DTMTKWECKNPYILSTKDYCPPRQELCLGNIDRIYDKNLMIKEHILA 518
QY 70 EAFIKAAAEFTLLROYNSKNVEDKILHRDITPEFRSMYTGCDRIDCLDIDISE 129
DB 519 IAIYES---RIUKRTRYKKDKDEVCITNK-----TFADIRDIIGGTDYWN 561
QY 130 KIADHDVTTAKKKTITAVFQKIGSKTNGKRV-----LEREGMKKEGLSTWKMGLCA 181
DB 562 DLSNRKL-----VGKINTNSKYVHRNKKNDKLFREWMKVKYIKDWMN----- 603
QY 182 LSTINETKMDGCVTRYLTKYIKNDI-KEYIEEFPASRPPLRWVTEGDEDFVKNRKE 240
DB 604 -----VISWFKDKTYCKE--DIEENIPQEFRRFSEMGDDYCDQKTRM 644
QY 241 LVSLKKKDCSCTLRNNGTSNKTCDDNENGACCKTOCEKRYKKMERKKHYSSQKKFOLY 300
DB 645 IETLAKEC-----KEKPEDDN---CKSKCNSTKEMVSKKKEEKNQAKOYOEY 690
QY 301 KNSATYNGILAVKEANSEFY--KNDEPVTSEANSKAKHARDYLKTOLENMICT 349
DB 691 QKGNNTKMYSEFKSIKPEYVLLKYSKCSNLNFEDEFKEIHSIDYKMK-CT 740

RESULT 15
US-08-568-459A-17
Sequence 17, Application US/08568459A
Patent No. 5849306
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568,459A
FILING DATE: 07-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 324 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-568-459A-17

Query Match
Best local Similarity 14.8%; Score 284; DB 2; Length 324;
Matches 74; Conservative 15; Mismatches 145; Indels 18; Gaps 6;

QY 43 CMPPRONLCVHYL--TKLNDSEKEDLRFAFKSAAEFTLLROYNSKNVEDKILHR 100
DB 25 CIPPRQKLCIHLVTLXXXXXXKXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 84
QY 101 DMIPPEFRSMETPEDYRDICLDTDISKILADHDVTTAKKTTAVFQIGSTTNGKV 160
DB 85 XXXXXDFKQMFYTFADYRDICLGTDISK---KDTXXXXXXKXXXXXXXXXISN--SI 139
QY 161 LEREGMKKEGLSTWKMGLCAL-----SYNFTETKMDGCVTRYLTKYIKNDIKEYLE 214
DB 140 RYKSKWETNGPIYTWGMLCALXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 199
QY 215 EFASRPPLRWVTEGDEDFVKNRKKELVSLKKKDCSCTLRNNGTSNKTCDDNENGACKT 274
DB 200 XXXXRPQFLRWLTWEGENFCCKEQKKEYVLLAK---CXXXXXXKXXXXXC--XXXXVACND 254
QY 275 QCEKYKKMKMERW 286
DB 255 QCKYHSHWIGIW 266
```

Search completed: June 20, 2003, 15:07:02
Job time : 20.1745 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 15:05:30 ; Search time 36.7511 seconds
(without alignments)
1033.453 Million cell updates/sec

Title: US-10-087-013-9

Perfect score: 1913

Sequence: 1 PCKWQKLISEQIEKNININ.....AKHARDYLTQLENNICTNG 351

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Published_Applications_AA.*
2: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|-------------------|
| 1 | 646 | 33.8 | 921 | US-10-153-273-8 | Sequence 8, Appl |
| 2 | 625.5 | 32.7 | 2710 | US-10-153-273-12 | Sequence 12, Appl |
| 3 | 353 | 18.5 | 749 | US-10-153-273-6 | Sequence 6, Appl |
| 4 | 353 | 18.5 | 1086 | US-09-924-154-15 | Sequence 15, Appl |
| 5 | 333.5 | 17.4 | 1143 | US-09-924-154-14 | Sequence 14, Appl |
| 6 | 329 | 17.2 | 1135 | US-10-153-273-4 | Sequence 4, Appl |
| 7 | 312 | 16.3 | 1421 | US-09-924-154-13 | Sequence 13, Appl |
| 8 | 284 | 14.8 | 324 | US-10-153-273-17 | Sequence 17, Appl |
| 9 | 272.5 | 14.2 | 311 | US-10-153-273-21 | Sequence 21, Appl |
| 10 | 269 | 14.1 | 1115 | US-10-153-273-2 | Sequence 2, Appl |
| 11 | 231 | 12.1 | 700 | US-10-153-273-10 | Sequence 10, Appl |
| 12 | 226.5 | 11.8 | 972 | US-09-924-154-16 | Sequence 16, Appl |
| 13 | 220 | 11.5 | 1501 | US-09-924-154-17 | Sequence 17, Appl |
| 14 | 171 | 8.9 | 277 | US-10-153-273-15 | Sequence 15, Appl |
| 15 | 144.5 | 7.6 | 411 | US-10-153-273-20 | Sequence 20, Appl |
| 16 | 137 | 7.2 | 282 | US-10-153-273-16 | Sequence 16, Appl |
| 17 | 135 | 7.1 | 291 | US-10-153-273-13 | Sequence 13, Appl |
| 18 | 131.5 | 6.9 | 362 | US-10-153-273-18 | Sequence 18, Appl |
| 19 | 101 | 5.3 | 1167 | US-09-815-242-11522 | Sequence 11522, A |

| | | | | | | |
|----|------|-----|------|----|---------------------|-------------------|
| 20 | 101 | 5.3 | 1938 | 9 | US-10-171-311-164 | Sequence 164, App |
| 21 | 101 | 5.3 | 1945 | 9 | US-09-927-597-2 | Sequence 2, Appl |
| 22 | 1972 | 5.3 | 1972 | 9 | US-10-171-311-162 | Sequence 162, App |
| 23 | 101 | 5.3 | 1979 | 9 | US-09-927-597-4 | Sequence 4, Appl |
| 24 | 99 | 5.2 | 874 | 9 | US-09-893-519A-50 | Sequence 50, Appl |
| 25 | 98.5 | 5.1 | 1092 | 9 | US-09-423-126-5 | Sequence 5, Appl |
| 26 | 98 | 5.1 | 380 | 10 | US-09-134-333-13 | Sequence 13, Appl |
| 27 | 98 | 5.1 | 456 | 10 | US-09-925-300-1595 | Sequence 1595, Ap |
| 28 | 98 | 5.1 | 1805 | 9 | US-09-820-843A-73 | Sequence 73, Appl |
| 29 | 97.5 | 5.1 | 665 | 9 | US-09-820-843A-107 | Sequence 107, App |
| 30 | 97 | 5.1 | 208 | 9 | US-10-084-205-52 | Sequence 52, Appl |
| 31 | 97 | 5.1 | 208 | 10 | US-09-925-637-52 | Sequence 52, Appl |
| 32 | 97 | 5.1 | 411 | 9 | US-10-153-273-19 | Sequence 19, Appl |
| 33 | 97 | 5.1 | 711 | 9 | US-09-298-523B-3 | Sequence 3, Appl |
| 34 | 95 | 5.0 | 677 | 10 | US-09-745-763-168 | Sequence 168, App |
| 35 | 95 | 5.0 | 716 | 10 | US-09-815-242-12403 | Sequence 12403, A |
| 36 | 95 | 5.0 | 1475 | 10 | US-09-740-274-2 | Sequence 2, Appl |
| 37 | 94.5 | 4.9 | 271 | 9 | US-10-153-273-14 | Sequence 14, Appl |
| 38 | 94.5 | 4.9 | 2353 | 10 | US-09-797-862-33 | Sequence 33, Appl |
| 39 | 93 | 4.9 | 380 | 10 | US-09-134-333-12 | Sequence 12, Appl |
| 40 | 93 | 4.9 | 654 | 10 | US-09-952-013A-5 | Sequence 5, Appl |
| 41 | 91 | 4.8 | 315 | 9 | US-10-106-698-5557 | Sequence 5557, Ap |
| 42 | 90 | 4.7 | 610 | 9 | US-09-802-640-36 | Sequence 36, Appl |
| 43 | 90 | 4.7 | 621 | 9 | US-10-205-823-357 | Sequence 357, App |
| 44 | 89.5 | 4.7 | 621 | 10 | US-09-856-247A-2 | Sequence 2, Appl |
| 45 | 89.5 | 4.7 | 670 | 9 | US-09-298-523B-63 | Sequence 63, Appl |

ALIGNMENTS

RESULT 1
US-10-153-273-8
Sequence 8, Application US/10153273
Patent No. US20020169305A1
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
Chitnals, Chetan
Miller, Louis H.
Peterson, David S.
Su, Xin-zhuang
Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/153,273
FILING DATE: 21-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:

LENGTH: 921 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Plasmodium falciparum
 SEQUENCE DESCRIPTION: SEQ ID NO: 8:
 US-10-153-273-8

Query Match 33.8%; Score 646; DB 9; Length 921;
 Best Local Similarity 41.0%; Pred. No. 5,3e-45;
 Matches 150; Conservative 54; Mismatches 112; Indels 50; Gaps 16;

QY 2 CKMYOKLISEOLEKNNHNCCKTED-AKMCENT--KIGDEGYCMPPRONLCVHYLT 58
 DB 396 CEIYAEMLKDNKNGRTVGECEYRKETYSMTCDSEKIKMGQ-HGACIPPRRCKLHYLEK 454
 QY 59 LNDKSEEDLEAREFKSAAEFTLLROY--NSKNVED-DKILHRDIPPEFRSMFYT 114
 455 IMTWNNE--LKAFKCAAAETFLMQYKKDKNGMDELDEKLGKGIPEDFKROMFYT 512
 QY 115 FGDRDICTDPTDI-SEKIAHDVTTAKKTTAVQKIGSKTTNGKVLEREKWKEXGLS 173
 DB 513 FADYRDICLGDISSKDKTSKGVKVCNIDVFKYISN-----SIRYKRWETNGPV 566
 QY 174 IMKMGALASYNT-----ET-KKMDGVRTYLMKYIKNNDIKEYLEEFASRPPLRW 225
 DB 567 IMEMLCALSYDTSLNANVPETHKRLKEGNNNF-EKVIYF-GSDSSTLTSFSEKPOPLRW 624
 QY 226 VTEMGEDPVKNNRKKELVSLKKCDSCITLRNNGTSNKTCDNENCGACKTOCEKRYKKMER 285
 DB 625 LTEMGEFCEKQKKEYKVLAKCKCDVDGQ---KC--NGKVCACKDCKOYHSHIGI 678
 QY 286 WKKYSOKKKFOLYKKSATYNNGLAVKEANSEYTKNDPEVTEANSKHAHDYLTQLEN 345
 DB 679 WIDWYKOKGKRYTEVK-----KIPLYKEDKDVKNSSD---ARDYLKTQLEN 721
 QY 346 MICTNG 351
 DB 722 MKCVNG 727

RESULT 2
 US-10-153-273-12
 ; Sequence 12, Application US/10153273
 ; Patent No. US20020169305A1
 GENERAL INFORMATION:
 APPLICANT: Sim, Kim L.
 Chitnis, Chetan
 Miller, Louis H.
 Peterson, David S.
 Su, Xin-zhaun
 Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/153,273
 FILING DATE: 21-May-2002

CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/210,288
 FILING DATE: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Fuller, Michael
 REGISTRATION NUMBER: 36,516
 REFERENCE/DOCKET NUMBER: NIH121,1FMDV1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2710 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Plasmodium falciparum
 SEQUENCE DESCRIPTION: SEQ ID NO: 12:
 US-10-153-273-12

Query Match 32.7%; Score 625.5; DB 9; Length 2710;
 Best Local Similarity 38.8%; Pred. No. 1.1e-42;
 Matches 137; Conservative 61; Mismatches 126; Indels 29; Gaps 9;

QY 2 CKMYOKLISEOLEKNNHNCCKTED-AKMCENTKIGDEGYCMPPRONLCVHYLT--T 57
 DB 1349 CKIVERILEBKNNRTVGEENPREKSTPMDCKNNIDISHDGACMPRRRCKLIYIAHSS 1408
 QY 58 KLNDKSEEDLEAREFKSAAEFTLLROYNNKNNVEDDKILHRDIPPEFRSMFYTFD 117
 DB 1409 QTEINIKTDNLKAPFTAAETFLSMOYKSKNDSEAKILDRGLIPSOFLSMYTFED 1468
 QY 118 YRDICLDTDISEKIAHDVTTAKKTTAVFORIGSKTTNGKVLEREKWKEXGLSITWG 177
 DB 1469 YRDICLNTDISK--QNDYAKAKADKIGKFSKDGSKSPSG---LSRQEWKNGPEIWK 1523
 QY 178 MCAIS--YNTETKKMDGVRTYLMKYIKNNDIKEYLEEFASRPPLRWTEMGEDPVKN 236
 DB 1524 MCALTRYVTDTDNKRKIKNDYSYDKVNSQNGNPSLEFFAKPOFLRMIMTEGEEFCE 1583
 QY 237 RKKEVLSLKKCDSCITLRNNGTSNKTCDNENCGACKTOCEKRYKKMERKKNYSQKK 296
 DB 1584 ROKKENTITDQAC-----NEINSTQCCNDAKH--RCNACRAYOEVYENKKKEFGQTNN 1635
 QY 297 FOLYKKSATYNNGLAVKEANSEYTKNDPEVTEANSKHAHDYLTQLENMICT 349
 DB 1636 FYLKANVQPD-----PEYKGYEYKDGVPQIGN-----EYLLOKCDNNKCS 1677

RESULT 3
 US-10-153-273-6
 ; Sequence 6, Application US/10153273
 ; Patent No. US20020169305A1
 GENERAL INFORMATION:
 APPLICANT: Sim, Kim L.
 Chitnis, Chetan
 Miller, Louis H.
 Peterson, David S.
 Su, Xin-zhaun
 Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US

```

ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/153,273
FILING DATE: 21-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH21,1FMDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 749 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-153-273-6

```

```

Query Match      18.5% Score 353; DB 9; Length 749;
Best Local Similarity 28.2% Pred. No. 6,8e-21;
Matches 103; Conservative 50; Mismatches 118; Indels 94; Gaps 16;

```

```

QY 10 ISEIERNINHC-KTEDAKKWC-ENTKLGDEGVCMPPRRONLCVHYLTKLNDOSKE 65
    103 SKHESSVTCCTKTSKYKKKNCYSNNKYTPREGVCGPPRQOCLGTFILIR-DGNE 161
    66 EDLEAFIKSAAEFTLLROYNSKNVEDDKILHRDIMPPEFRSMFTYRGDYRDLCLDT 125
    162 EGLKDHINKAANYEAMHLKEKY--ENAGGDKICN-----AIIGSYADIGDIVRGL 209
    126 DISEKADHDVYTAKKITVAVFOKIGSKTTNGKKVL-----EREGWKEYLSTWKGMLCA 181
    210 DVA-----RDINT--NKLSEKFOKIFMGGSNRKKONDNNERNKRWKMKOHNLIWSSM-- 259
    182 LSYNTEFKKMDGVRTYLMKYIKNNNDIKYELEEFASRPPLRWVTWGEDFVKNRKKEL 241
    260 -----VKHTRKGTCKRH--NNEKIFPFLRWLWKEMGDFECDEMGTEV 300
    QY 242 VSLKKKDCSTLRNNGTSNKTCDNENCGACKTQCEKRYKKWEMRKKHYSSOKKKF--- 297
    DB 301 KOLEKICE-----NKNCSSEK---CKNACSSYEKWKERKNEYMLQSKKFPDSK 346
    QY 298 QLYKNSATYNNGLAVKEANSETYK-----NDPEVTANSKKHARDYLKTOLEN 345
    DB 347 KLKKKNNLVN-----KFEDSKAYLRSESKOCNIEFDEFTFPNKYKEA-----C 392
    QY 346 MICTN 350
    DB 393 MGCEN 397

```

```

RESULT 4
US-09-924-154-15
Sequence 15, Application US/09924154
Patent No. US20020127241A1
GENERAL INFORMATION:
APPLICANT: Natum, David L.
APPLICANT: Sim, Kim L.

```

```

TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
FILE REFERENCE: 05213-0465 43170-262105
CURRENT APPLICATION NUMBER: US/09/924,154
CURRENT FILING DATE: 2001-08-07
PRIOR APPLICATION NUMBER: US 60/223,525
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin version 3.1
SEQ ID NO 15
LENGTH: 1086
TYPE: PRT
ORGANISM: Mammalian
US-09-924-154-15

```

```

Query Match      18.5% Score 353; DB 10; Length 1086;
Best Local Similarity 28.2% Pred. No. 1,1e-20;
Matches 103; Conservative 50; Mismatches 118; Indels 94; Gaps 16;

```

```

QY 10 ISEIERNINHC-KTEDAKKWC-ENTKLGDEGVCMPPRRONLCVHYLTKLNDOSKE 65
    DB 345 SKHESSVTCCTKTSKYKKKNCYSNNKYTPREGVCGPPRQOCLGTFILIR-DGNE 403
    66 EDLEAFIKSAAEFTLLROYNSKNVEDDKILHRDIMPPEFRSMFTYRGDYRDLCLDT 125
    DB 404 EGLKDHINKAANYEAMHLKEKY--ENAGGDKICN-----AIIGSYADIGDIVRGL 451
    QY 126 DISEKADHDVYTAKKITVAVFOKIGSKTTNGKKVL-----EREGWKEYLSTWKGMLCA 181
    DB 452 DVA-----RDINT--NKLSEKFOKIFMGGSNRKKONDNNERNKRWKMKOHNLIWSSM-- 501
    QY 182 LSYNTEFKKMDGVRTYLMKYIKNNNDIKYELEEFASRPPLRWVTWGEDFVKNRKKEL 241
    DB 502 -----VKHTRKGTCKRH--NNEKIFPFLRWLWKEMGDFECDEMGTEV 542
    QY 242 VSLKKKDCSTLRNNGTSNKTCDNENCGACKTQCEKRYKKWEMRKKHYSSOKKKF--- 297
    DB 543 KOLEKICE-----NKNCSSEK---CKNACSSYEKWKERKNEYMLQSKKFPDSK 588
    QY 298 QLYKNSATYNNGLAVKEANSETYK-----NDPEVTANSKKHARDYLKTOLEN 345
    DB 589 KLKKKNNLVN-----KFEDSKAYLRSESKOCNIEFDEFTFPNKYKEA-----C 634
    QY 346 MICTN 350
    DB 635 MGCEN 639

```

```

RESULT 5
US-09-924-154-14
Sequence 14, Application US/09924154
Patent No. US20020127241A1
GENERAL INFORMATION:
APPLICANT: Natum, David L.
APPLICANT: Sim, Kim L.
TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
FILE REFERENCE: 05213-0465 43170-262105
CURRENT APPLICATION NUMBER: US/09/924,154
CURRENT FILING DATE: 2001-08-07
PRIOR APPLICATION NUMBER: US 60/223,525
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin version 3.1
SEQ ID NO 14
LENGTH: 1143
TYPE: PRT
ORGANISM: Mammalian
US-09-924-154-14

```

```

Query Match      17.4% Score 333.5; DB 10; Length 1143;
Best Local Similarity 27.9% Pred. No. 4,8e-19;
Matches 92; Conservative 46; Mismatches 131; Indels 61; Gaps 11;

```

```

QY 9 ISEIERNINH---NCKTEDAKKWCENTKLGDEGVCMPPRRONLCVHYLTKLNDOSK 64

```

Db 448 IASQINVNDLRGCGCNKSNNEKSMWCTGTFNKPCTCEPPRQRLCLG-RTYLLHRGH 506
65 EEDLEAFIKSAAEFTLLROYNSKNVEDDKILHRMIPPEFFRSMEFTYEGDRLDCLD 124
507 EEDYKELLASIIYEAOULLKYKKE--DENAL-----CSTIÖNSYADLADIKG 554
125 TDISEKADHDVTTAKKITAVFQKIGSKTTNGKYLE--REGWMEYGLSIMGMLCAL 182
555 SDI---IKDY-----YGKKMEENLNKYNKDKKRNESLKIFREKWMDEKNKENVK----- 601
183 SYNTETKKMDEGVRTYLMKTIYKNNDIKEYLEEFASRPPLRWTEGDEDFVKNRKELY 242
602 -----VMSAVLNKNETCKDYDKFQKIPQFLRMFKEMGDDFCERKREKIY 645
243 SLKKKDCSTLRNNGTSNKTCDNENCGACKTOCEKTKMMERWKHHYSOKKRFOLYKN 302
646 SFESFKVECC-----KKKCCDEN---TKNKCSEYKKWIDLKSEYKQVDKYTKDN 694
QY 303 SATYNN--GLAVEANSEYTKNDPEVTEAN 330
695 KKMVDNIDEVKNKEANYLLEKESKECKDVN 724

RESULT 6

US-10-153-273-4
Sequence 4, Application US/10153273
Patent No. US20020169305A1

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

Chitnis, Chetan

Miller, Louis H.

Peterson, David S.

Su, Xin-zhuan

Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/153,273

FILING DATE: 21-May-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/210,288

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Fuller, Michael

REGISTRATION NUMBER: 36,516

REFERENCE/DOCKET NUMBER: NIH21.1FMDV1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1435 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHEICAL: NO

ORIGINAL SOURCE:

ORGANISM: Plasmodium falciparum
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-153-273-4

Query Match 17.28; Score 329; DB 9; Length 1435;
Best Local Similarity 24.88; Pred. No. 1.5e-18;
Matches 87; Conservative 63; Mismatches 113; Indels 88; Gaps 13;

QY 14 EKNNHCKTTEADAKKCEWTKLGEDEGVCMPPRRQALCVHYLTKNLDS-----KEEDLR 69
Db 463 DKNSV-----DTNTRWCKKNPILSTKDVCPVPRQELCIGNIDRYDKMLMIKEHILA 518
QY 70 EAFIKSAAEFTLLROYNSKNVEDDKILHRMIPPEFFRSMEFTYEGDYDIDCLDISE 129
Db 519 IATYES-----RLKRRKKNNDKDEVCITIK-----TFADIDIDIGTDYWN 561
QY 130 KIADHDVTTAKKITAVFQKIGSKTTNGKVV-----LERGWMKEYGLSIMGMLCA 181
Db 562 DLSNRKL-----VGRINTSKYVHNRKKNKDLFRDEWVKVIRKDVWN----- 603
QY 182 LSYNETKKMDEGVRTYLMKTIYKNNDI--KEYLEEFASRPPLRWTEGDEDFVKNRKE 240
Db 604 -----VISWVERDKTVCKE--DQIENIPQFFRSEWGDYCOQDKTKM 644
QY 241 LVSLKKKDCSTLRNNGTSNKTCDNENCGACKTOCEKTKMMERWKHHYSOKKRFOLY 300
Db 645 IETLVKVC-----KKPCEDDN---CKSKCNSEYKWTISKKKEEYNKQAKOQY 690
QY 301 KNSATYNNGLAVEANSEY--KNDEPVTANSAKHARDYLKTQLENMICT 349
Db 691 QKGNMYKMYSEFKSIKPEVYLKRYSEKSNLNEDEFKEELHSDYKNK--CT 740

RESULT 7

US-09-924-154-13
Sequence 13, Application US/09924154
Patent No. US20020127241A1

GENERAL INFORMATION:

APPLICANT: Narum, David L.

APPLICANT: Sim, Kim L.

TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use

FILE REFERENCE: 05213-0465 43170-262105

CURRENT APPLICATION NUMBER: US/09/924,154

CURRENT FILING DATE: 2001-08-07

PRIOR APPLICATION NUMBER: US 60/223,525

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 17

SOFTWARE: Patentin version 3.1

SEQ ID NO 13

TYPE: PRT

ORGANISM: Mammalian

US-09-924-154-13

Query Match 16.38; Score 312; DB 10; Length 1421;
Best Local Similarity 24.28; Pred. No. 3.8e-17;
Matches 85; Conservative 61; Mismatches 117; Indels 88; Gaps 13;

QY 14 EKNNHCKTTEADAKKCEWTKLGEDEGVCMPPRRQALCVHYLTKNLDS-----KEEDLR 69
Db 463 DKNSV-----DTNTRWCKKNPILSTKDVCPVPRQELCIGNIDRYDKMLMIKEHILA 518
QY 70 EAFIKSAAEFTLLROYNSKNVEDDKILHRMIPPEFFRSMEFTYEGDYDIDCLDISE 129
Db 519 IATYES-----RLKRRKKNNDKDEVCITIK-----TFADIDIDIGTDYWN 561
QY 130 KIADHDVTTAKKITAVFQKIGSKTTNGKVV-----LERGWMKEYGLSIMGMLCA 181
Db 562 DLSNRKL-----VGRINTSKYVHNRKKNKDLFRDEWVKVIRKDVWN----- 603
QY 182 LSYNETKKMDEGVRTYLMKTIYKNNDI--KEYLEEFASRPPLRWTEGDEDFVKNRKE 240
Db 604 -----VISWVERDKTVCKE--DQIENIPQFFRSEWGDYCOQDKTKM 644


```

Db      XXXXXDFKQMFYTFADYDIDIGDTISKK---KOTISXXXXXXXXXXXXXSI--SI 139
QY      85
Db      161 LEDEGWMKEGLSINKGMICAL-----SYNETKMGDCGVRTYLKITYKNNDKEYIE 214
          : : : : :
Db      140 RYKSMWETNGPIVIMGMLCALXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 199
          : : : : :
QY      215 EFARSPPEFLRMVEMEDVEFKNRKRELYSLKKKCDSCTLRNNGTSNKTCDGNCGACT 274
          |||||:||||:|:||| | | | |
Db      200 XXXXRQPLRLWLFEMENECBEQRKYVLAK---CXXXXXXXXXXC-XXXVCACRD 254
          : : : : :
QY      275 QCEYRKMMERW 286
          ||::| | : |
Db      255 OCKOYHSWIGIW 266

RESULT 9
US-10-153-273-21
; Sequence 21, Application US/10153273
; Patent No. US20020169305A1
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
           Chitnis, Chetan
           Miller, Louis H.
           Peterson, David S.
           Su, Xin-zhaun
           Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
                   AND PLASMODIUM FALCIPARUM ERITHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
City: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/153,273
FILING DATE: 21-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH21.1FMV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 311 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEtical: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-153-273-21

Query Match      14 28; Score 272.5; DB 9; Length 311;
Best Local Similarity 28.48; Pred. No. 1e-14;
Matches 71; Conservative 22; Mismatches 140; Indels 17; Gaps 6;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 14:59:06 ; Search time 27.6965 Seconds

(without alignments)
1218.319 Million cell updates/sec

Title: US-10-087-013-9

Perfect score: 1913

Sequence: 1 PKMWOKLISEQIEKNINHN.....AKHARDYLKTOLENNICTNG 351

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

1 number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 625.5 | 32.7 | 3078 | 2 T28432 | variant-specific s |
| 2 | 562.5 | 29.4 | 3026 | 2 T28431 | variant surface pr |
| 3 | 547.5 | 28.6 | 1711 | 2 C71625 | variant-specific s |
| 4 | 520 | 27.2 | 2706 | 2 T28155 | variant-specific s |
| 5 | 501 | 26.2 | 3006 | 2 T28625 | variant-specific s |
| 6 | 443 | 23.2 | 2212 | 2 T28157 | erythrocyte membra |
| 7 | 433.5 | 22.7 | 2042 | 2 T18399 | variant-specific s |
| 8 | 432 | 22.6 | 2924 | 2 T18378 | variant-specific s |
| 9 | 403.5 | 21.1 | 2647 | 2 T28161 | hypothetical prote |
| 10 | 399.5 | 20.9 | 2664 | 2 T28626 | variant-specific s |
| 11 | 398.5 | 20.8 | 2197 | 2 B71600 | variant-specific s |
| 12 | 397.5 | 20.8 | 2135 | 2 T14602 | variant-specific s |
| 13 | 375 | 19.6 | 2228 | 2 T14029 | variant-specific s |
| 14 | 374.5 | 19.6 | 2182 | 2 T28634 | variant-specific s |
| 15 | 360.5 | 18.8 | 1729 | 2 T18396 | erythrocyte membra |
| 16 | 329 | 17.2 | 1435 | 2 A37793 | erythrocyte-bindin |
| 17 | 269 | 14.1 | 1070 | 2 T30848 | Duffy receptor - p |
| 18 | 261 | 13.6 | 1153 | 2 T28652 | erythrocyte bindin |
| 19 | 254 | 13.3 | 1045 | 2 T18373 | erythrocyte bindin |
| 20 | 171.5 | 9.0 | 778 | 2 A35970 | erythrocyte bindin |
| 21 | 127.5 | 6.7 | 1123 | 2 T28139 | transforming prote |
| 22 | 122.5 | 6.4 | 690 | 2 T51298 | asparagine/asparta |
| 23 | 118 | 6.2 | 3844 | 2 T18402 | hypothetical prote |
| 24 | 117.5 | 6.1 | 911 | 2 S51441 | major surface glyco |
| 25 | 117.5 | 6.1 | 1004 | 2 T28221 | hypothetical prote |
| 26 | 116.5 | 6.1 | 335 | 2 T25110 | protein T22C1.6 [1 |
| 27 | 116.5 | 6.1 | 346 | 2 F87844 | thopyry protein - |
| 28 | 115.5 | 6.0 | 2401 | 2 T28676 | spindle pole body |
| 29 | 115 | 6.0 | 944 | 2 S26710 | |

| | | | | | |
|----|-------|-----|------|-----------|---------------------|
| 30 | 114 | 6.0 | 981 | 2 C82930 | p115 protein U0140 |
| 31 | 114 | 6.0 | 993 | 2 S49461 | synaptonemal compl |
| 32 | 114 | 6.0 | 2269 | 2 T28677 | thopyry protein |
| 33 | 113 | 5.9 | 1390 | 2 S51364 | sperm tail-specific |
| 34 | 111.5 | 5.8 | 837 | 2 T48407 | hypothetical prote |
| 35 | 111.5 | 5.8 | 1939 | 2 T18372 | repeat organellar |
| 36 | 110.5 | 5.8 | 1346 | 2 G71613 | hypothetical prote |
| 37 | 110.5 | 5.8 | 1712 | 2 C71618 | hypothetical prote |
| 38 | 110 | 5.8 | 1995 | 2 T08166 | probable membrane |
| 39 | 109 | 5.7 | 482 | 2 T24518 | hypothetical prote |
| 40 | 109 | 5.7 | 845 | 2 T148176 | synaptonemal compl |
| 41 | 109 | 5.7 | 2139 | 2 T18296 | myosin heavy chain |
| 42 | 108 | 5.6 | 508 | 2 E71620 | hypothetical prote |
| 43 | 108 | 5.6 | 3335 | 2 H81702 | adherence factor T |
| 44 | 107.5 | 5.6 | 2910 | 2 T28156 | DNA-directed RNA p |
| 45 | 107 | 5.6 | 657 | 2 G71602 | protein with DnaJ |

ALIGNMENTS

RESULT 1

T28432

variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)

N:Alternate names: erythrocyte membrane binding protein 1 (EMP1)

C:Species: Plasmodium falciparum

C>Date: 15-Oct-1999 #sequence-revision 15-Oct-1999 #text-change 09-Jun-2000

C/Accession: T28432

R:Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herfelde, J.A.; Peterson, Cell 82, 89-100, 1995

A>Title: The large diverse gene family var encodes proteins involved in cytoadherence

A:Reference number: Z20487, PMID:95330813, PMID:7606788

A:Accession: T28432

A:Molecule type: DNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-3078 <SU>

A:Cross-references: EMBL:L40608; NID:g886374; PID:g886375; PIDN:AA75396.1

C:Genetics:

A:Gene: var-1

A:Introns: 261/3

| Query Match | Score | DB 2; | Length | 3078; |
|---|------------------|---|------------|---------|
| Best Local Similarity | 38.8% | Pred. No. 4.7e-35; | | |
| Matches 137; | Conservative 61; | Mismatches 126; | Indels 29; | Gaps 9; |
| QY | 2 | CKMWOKLISEQIEKNINHNCKTED-AKWCENTKLGDEGVCMPPRONLCVHYL--T 57 | | |
| DB | 1347 | CKIVEKILEGKNGRTTVGECNPKESYPDMCKNNIDISHGACMPPRORCLCYIAHES 1406 | | |
| QY | 58 | KLNDDSKEDLREAFIKSAAEFFLLROYNSKNVEDDKTLHRDMTPREFFRSMFTFGD 117 | | |
| DB | 1407 | QENIKITDMLKDAEFTKTAAEFFLLSMOYKSKNSDEAKLIDGLIPSOFLRSMFTFGD 1466 | | |
| QY | 118 | YRDICLDTISEKIAHDVTTAKKITYAVFOKIGSTTNGKYLLEEGWAKKEGLSIWK 177 | | |
| DB | 1467 | YRICLINTDLSK--QNDVAKAKDKIGKFFSKSGSPSS--LSQDEWMTKTGPELWK 1521 | | |
| QY | 178 | MCALS-YNTEYKKMDGYRTYLMKYIYKNDIKLEYLEEFASPPPLRWVTEGDEPVKN 236 | | |
| DB | 1522 | MCALTKYVYDTPDKKRKKINDYSYDVKNOSQNPISLEFAAKPOFLRWMIWEGEEFCAE 1581 | | |
| QY | 237 | RKKELVSLKKKCCSCLRNNGTSNKTCDDENGACKTOCEYKKMMERKKHYSQKK 296 | | |
| DB | 1582 | RKKENITKDAC-----NEINSTOOCNDK--RCNQACRAAYEVENKKKFEFSQNTN 1633 | | |
| QY | 297 | FOLYKNSATYNNGLAVKANSEYTKNDPEVTEANSKHAHDYLTLENNICT 349 | | |
| DB | 1634 | FVLKANVQPD-----PEYKGYEKDGVQIQGN-----EYLLQCDNNKCS 1675 | | |
| RESULT 2 | | | | |
| T28431 | | | | |
| variant surface protein 1 homolog A4VAR - malaria parasite (Plasmodium falciparum) (f | | | | |

| | | | | |
|---------------------------|--------|--------------------|------------|--------------|
| Query Match | 98.6% | Score 547.5; | DB 2; | Length 1711; |
| Best Local Similarity | 37.3%; | Pred. No. 6.2e-30; | | |
| Matches 188; Conservative | 53; | Mismatches 105; | Indels 57; | Gaps 14; |

282 WMRKKKHYSÖKKKÖLYKKSATYNN---GLAVE--ANSERYKKNDE---VTEANS 331

OY 332 AKH-ARDYLKTOLE 344
 Db 1097 AONSATEEYKKNIE 1110

RESULT 5

variant-specific surface protein 3 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T28625
 R:Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herfelde, J.A.; Peterson, D.S.
 Cell 82, 89-100, 1995

A:Title: The large diverse gene family var encodes proteins involved in cytoadherence at
 A:Reference number: Z20487; MID:95330813; PMID:7606788

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Accession: T28625
 A:Molecule type: DNA

A:Residues: 1-3006 <SU>

A:Cross-references: EMBL:LA0609; MID:9886376; PID:9886377; PIDN:AAV5397.1

C:Genetics:

A:Gene: var-3

A:Insertions: 2597/3

Query Match 26.2%; Score 501; DB 2; Length 3006;

Best Local Similarity 30.0%; Pred. No. 2e-26;

Matches 125; Conservative 63; Mismatches 123; Indels 106; Gaps 14;

OY 1 PCKMVKL-----ISEQIEKNINHNCKTEDAKMKCENTKLGEDGVCMPRRONLCVH 54

Db 1755 PCIVKTLFTTTELEACPTKYNGREKPPN--WKCISSGSDASGICIPRRRLYLH 1812

OY 55 YL-----TKLNDSEK-----EEDLRAPFKSAAEFPL-RQYNSKNVE----- 93

Db 1813 KIEGVDTVSSDDETTTPITHDALREAFIQTAAVEFFELHWRKIKIKERQELONGTF 1872

OY 94 -----DDKILHRDMIPEFFRSMFTFGDYRDIC-----LDTDL 127

Db 1873 LRPAPQKVPEDNPEHPQKKLKEGKIPPEFRKQMTLIGDYRDLGVKQVADVAQALEASG 1932

OY 128 SEKIADHDVTTAKKITTAVFQKIGSKTTNGK---VLEREGWKEVGLSIKMGICALS 183

Db 1933 DNKSGDKNIIKIDISEKISVLEKSGEQTTPGPKPGQTTTKPEEWQKGEHIMNAMICALT 1992

OY 184 YNTEKKNMGVATYTL-----KITYKNNDIKEY-----LEEFASRPF 222

Db 1993 HNTDTRQVDQVKGOLFENGNKTPKNSQYQKNTVITSSVSSNGPIGNIKLEQFASRPTF 2052

OY 223 LRWVTEGDEDFVNRKKELVSLKKKDCSTLRNNGTSNKTCDNE----- 267

Db 2053 LRHLEMGEEFCCKQYHKKLEKEEC-----HKDSNRKCCDDGFECKECPCKDGSFE 2105

OY 268 --NCGACKTCQCEKKYKMMERKMKRHSYSSQKKFOLYKNSATYNNGLAVKEANSETYKN 322

Db 2106 TLKCPSCAKSKCYKWKISRKDEFTKQGAVEKQKADAGNNN-DYKEF-SKTLRN 2160

OY 2106 TLKCPSCAKSKCYKWKISRKDEFTKQGAVEKQKADAGNNN-DYKEF-SKTLRN 2160

OY 2106 TLKCPSCAKSKCYKWKISRKDEFTKQGAVEKQKADAGNNN-DYKEF-SKTLRN 2160

OY 2106 TLKCPSCAKSKCYKWKISRKDEFTKQGAVEKQKADAGNNN-DYKEF-SKTLRN 2160

OY 2106 TLKCPSCAKSKCYKWKISRKDEFTKQGAVEKQKADAGNNN-DYKEF-SKTLRN 2160

OY 2106 TLKCPSCAKSKCYKWKISRKDEFTKQGAVEKQKADAGNNN-DYKEF-SKTLRN 2160

OY 2106 TLKCPSCAKSKCYKWKISRKDEFTKQGAVEKQKADAGNNN-DYKEF-SKTLRN 2160

OY 2106 TLKCPSCAKSKCYKWKISRKDEFTKQGAVEKQKADAGNNN-DYKEF-SKTLRN 2160

OY 2106 TLKCPSCAKSKCYKWKISRKDEFTKQGAVEKQKADAGNNN-DYKEF-SKTLRN 2160

OY 2106 TLKCPSCAKSKCYKWKISRKDEFTKQGAVEKQKADAGNNN-DYKEF-SKTLRN 2160

OY 2106 TLKCPSCAKSKCYKWKISRKDEFTKQGAVEKQKADAGNNN-DYKEF-SKTLRN 2160

C:Genetics:

A:Gene: EMP1

Query Match 23.2%; Score 443; DB 2; Length 2212;

Best Local Similarity 26.2%; Pred. No. 1.5e-22;

Matches 128; Conservative 59; Mismatches 133; Indels 168; Gaps 19;

OY 8 LISEQIEKNINHNCKTE-----DAKMKC-----ENTKLGED----- 39

Db 1272 IVAGVLTREKLEACPTKYNGREKPPN--WKCISSGSDASGICIPRRRLYLH 1331

OY 40 -----EGVCMPPRRONLCVHLYTKLND-----DSKEE 66

Db 1332 GSPVTSNSGICIPRRRLR---YIOLKHDWAGNTVSGAQTPOGTSPPSGKETPSD 1388

OY 67 DLREAFIKSAAEFPL-RQYNSKNVE-----DD 95

Db 1389 KLRTAFIQSAIETFFELMDRYKKEKEKEKRVANGLVPSLNGGPPQPGVYGDSPS 1448

OY 96 KILHRDMIPEFFRSMFTFGDYRDICL--DPTDISKILADHDVTTAKKITTAVFQKI-- 150

Db 1449 KLOQTGVIPPEFLRQMTLIGDYADIFFGKNDIYIDTKNGDKDIAREKKIKAIIEVYLK 1508

OY 151 --GSKTTNGKRVLEREGWKEVGLSIKMGICALSYNTETK-----KMDEGVATYTL-- 199

Db 1509 NADSPPEDEK---RQTMWQNGEHIWNGICALTYKEKDEKGPLKQNGLSALMDEK 1565

OY 200 -----MKTYL-----KNND--IKELYEFASRPPPLRWVTEGDFYKNRK 238

Db 1566 NKRPKDQKQYDVKLIDENSCTSPKTNHDVPPPLTNFISRPYFYLEWGEFFECREK 1625

OY 239 KELVSLKKKDCSTLRNNGTSNKTCDNEN-----CGACKTCQCEKKYK 282

Db 1626 KRLKIKVEK-----MDEGKKQKCSGDEGCEELRKQDYSTVNDFYCPREGKTCRETKRW 1681

OY 283 MEKWKHYSQKKFOLYKNSATYNNGLAVKEANSETYKNDPEVTEANSKAHARDYLKTO 342

Db 1682 IEKKKDEYDKQKRAYNNQKTDARNN-----NDNAFS-----TTLDCTTAGDPLQT- 1728

OY 343 LENNICTN 350

Db 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

OY 1729 LKNGPCKN 1736

```

Db      755 DKDGAICPPRRKRLYLHIEGV-DTDDKSLRKMFIESAAVETFLMDRKYKLTTPQSG 813
QY      90 -----KNED-----DKILHRMIPPEEFSRMFTGVDYRDICLDPDISEKA 132
Db      814 SPLLGWITGVGVENDDENNPEKILQKEIPDLROMFYTLIGDVRDLDFSGDKKKG 873
QY      133 DHDVYTA-----KKITAVFOKIGSKTTNGK-----VLERGWMKEYGLSIWKM 178
Db      874 YSDIVSGDNVIRENTIEKIASPFQNGNKGSTPHVKNPVOPTQTMKDNAKHIMWM 933
QY      179 LCAISYNTET-----KKMDEGVRTY-----LKKIYKND 208
Db      934 ICALTYEETSGSGDEKIEKDAYKAKFFGPNGNPLPQPTNGTSNPEISQYQDYV 993
QY      209 IKE-----YLEEFASRPFLRWYEMWGEDVEFKNKKELVSLK 245
Db      994 LKEENNAGMSTSPKSTASPDNPTTLTQLFVLRPTLYRLEWQCNFCCKEKKRLAQY 1053
QY      246 KKCDSCITLRNNGTSKCTD-----DNENGACKTQCEKYKMMERWK 288
Db      1054 EDC-----RGN-----DKVCSGDEDEEVKQDYSKISNFCGCGRECKRYKNWIKTKRT 1105
QY      289 HYSQKRRFOLYKNS-ATYNNG-----LAVKEANSEY-----KN- 322
Db      1106 EFDEQKKAIVDRKDKYTKTENKGAESKHSISDQNVKMLGTDYASITESFLERKISGCSKNN 1165
QY      323 -----DPEVT--EANSKAHARDYLKQLENNICTNG 351
Db      1166 KDNQDGTINFKEPVDYTFKPADNCKPCSEF-KVCENGCOSG 1206

```

RESULT 8

```

T18378
variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
N:Alternate names: erythrocyte membrane binding protein 1 (EMP1)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T18378
R:Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschl, T.
Cell 82, 77-87, 1995
A:Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and
A:Reference number: 218925; MUID:95330812; PMID:7541722
A:Accession: T18378
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2924 <BAR>
A:Cross-references: EMBL:U27338; NID:914918; PID:914919; PIDN:AAB60251.1
C:Genetics:
A:Gene: EMP1
A:Introns: 2476/3

```

```

Query Match      22.6%; Score 432; DB 2; Length 2924;
Best Local Similarity 26.3%; Pred. No. 1,2e-21;
Matches 132; Conservative 59; Mismatches 134; Indels 176; Gaps 19;

QY      2 CKMWQTLISEQIENNNHNCCKTEDAKWC-----ENKRLGDEG-VC 43
Db      839 CSYDALKKGLDACLTKTKTAPTSWKCPISGNNTTSTTKPGAAGTSGKDTGSIC 898
QY      44 MPRRQMLCVHYLTKLND-----LROYNSKN-----VEDDKIL 98
Db      899 VPPRRRL--YVKKLMDMAGGETTEAKSOETSGGQKTPSGNESPSPEKLPQGTPTTK 955
QY      62 DSKEDLREAFISAAETFL---LROYNSKN-----VEDDKIL 98
Db      956 ETPESSILHAFVSPFLRRFLPMHFKFEQWKAQAGATGQQTITGLDGGGETTPDKL 1015
QY      99 HRDMPPEEFSRMFTGVDYRDICL-DTDL-----SEKIAHDVYTA-KKITAVFOKIGS 152
Db      1016 KTHGIPDFLRQMYTLIGDVRDLVHTSGKEDMQIMEALOKRIEQLPTSGS 1075
QY      153 -KTNCKVLE--REGWMKEYGLSIWKMICALSYNTETK-----MDEGVRTYLM 200

```

```

Db      1076 SPSPPVDTQTSVENVNPKTWMNENGGKIWEGWVCAIYTNIDTPSGIATQIQEVRTKL 1135
QY      201 KTIYKNDIKET-----LEEFASRPFLRWYEMWGED 232
Db      1136 DENSKNPKLPQYKYDQVKLDDTSPDKTGSVPVSGEKITPLDFTSRPPYFRYLEWGET 1195
QY      233 FVKNRKKELVSLKKKCDSCITLRNNGTSNKT-----DNE-----NCACAKTQ 275
Db      1196 FCKEKKRLKEKKEC-----RGDRTGHEHSGSDGYDCTRIDADNRKFDVINCDCIHQ 1250
QY      276 CEKYKMMERWKHYSSQKKRQ-----LYKNSATYNNGLAVKEANSEYTKNDPEVTAN 330
Db      1251 CRKYRKWIDIKDEYHKEQEKYQGEYDKLTQKXSSGDNCKOLEKH----- 1298
QY      331 SAKHARDYLKQLENNICTNG 351
Db      1299 -KSAVAFLK--ELKHCKNG 1314

```

RESULT 9

```

T28161
hypothetical protein FCR3-varT11-1 - malaria parasite (Plasmodium falciparum) (fragme
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T28161
R:Hernandez-Rivas, R.; Matfel, D.; Sterkers, Y.; Peterson, D.S.; Wellens, T.E.; Scher
Mol. Cell. Biol. 17, 604-611, 1997
A:Title: Expressed var genes are found in Plasmodium falciparum subtelomeric regions.
A:Reference number: 220483; MUID:97154495; PMID:9001213
A:Accession: T28161
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2647 <HER>
A:Cross-references: EMBL:U67959; NID:9179425; PID:91809295; PIDN:AAC47438.1
A:Experimental source: strain FC027/PNG
C:Genetics:
A:Introns: 2158/3
A>Note: FCR3-varT11-1

```

```

Query Match      21.1%; Score 403.5; DB 2; Length 2647;
Best Local Similarity 25.1%; Pred. No. 1e-19;
Matches 125; Conservative 60; Mismatches 131; Indels 183; Gaps 20;

QY      8 LISQIEKNNHNCCKTE-----DAKWC-----ENKRLGDE----- 39
Db      1273 IVAGVLKRNENMNCPTKRYGKAPTSMKCIPTKTNATSGSGSGNALQRAKRAVES 1332
QY      40 -----EGVCMPPRRQMLCVHYLTKLND-----DSKEE 66
Db      1333 GSPVTSNSGSLCIPRRRL--YIQKLHWASGNVVGQAQTPQCGTSSPSGKETPSD 1389
QY      67 DLREAFISAAEFLROY-----NSKNVDDKIL--HRD----- 101
Db      1390 KLRTAFQSAIEFFLMDRKKGKALAKREKKQWVDYSPSTADPHNPVSLVIANP 1449
QY      102 -----MIPPEFSRMFTGVDYRDICL--DTDISEKIDHDVYTAKKITAVFOKI-- 150
Db      1450 NYKTCVPIPPFLROMYTLIGDADIFFGANDIYIDTKNDKDIAREKKIKAIKAEVVK 1509
QY      151 --GSKTTNGKVLREBGMWKEYGLSIWKMICALSYNTETK-----KMDEGVRTY-- 199
Db      1510 MADQPSDEK--RQTMWQNGEHIWNGICALTYKEKDEKTPKQNGGLKSALMDEK 1566
QY      200 -----MKIYKNDIKLEYLEEA-----SRPFLRWY 227
Db      1567 NKAPKDKQYD--KVKLDENSGTSPKIVPAPKPTTTPPPSPPTSPSPSPRYRYLE 1622
QY      228 EMGEDVFNKRRKELVSLKKKCDSCITLRNNGTSNKTDDNEN-----CGA 271
Db      1623 EMAETFCRBRKRLKELKIVC-----MDEQSKKQKCGSDGDCGEIKQDYSTYRVDYCE 1678
QY      272 CKQCEKYKMMERWKHYSSQKKRQLYKNSATYNNGLAVKEANSEYTKNDPEVTAN 331

```


Db 1679 CGKCYCFYKRWIGKKKDEYKOEKAYNNOKTDARRNN-----NDNAFS-----TTLDLT 1726
QY 332 AKHARDYKTOLENNICTN 350
Db 1727 CTTAGDFLOT-LKNGPCKN 1744

RESULT 10

T28626
variant-specific surface protein 2 - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T28626
R:Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herfelde, J.A.; Peterson, D.S.
Cell 82, 89-100, 1995
A:Title: The large diverse gene family var encodes proteins involved in cytoadherence and
A:Reference number: Z20487; M0ID:95330813; PMID:7606788
A:Accession: T28626
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2664 <SD>
A:Cross-references: EMBL:L40609; NID:g886376; PID:g886378; PIDN:AAA75398.1
C:Genetics:
A:Introns: 2197/3
A:Note: var-2

Query Match 20.9%; Score 399.5; DB 2; Length 2664;
Best Local Similarity 25.2%; Pred. No. 1.9e-19;
Matches 124; Conservative 58; Mismatches 118; Indels 193; Gaps 21;

QY 1 PCKMVKLISEQIEKNNHCKTEBDK-----WKCENTK-----LGED----- 39
Db 1281 PCQIVKELFS-----NVDTLQACSTKYGPKKNYMRCPPTTSNDVIGEDGQSRV 1334
QY 40 -----EGYCMPPRONLCVHYLTKND----- 61
Db 1335 RSTPESGNSDKKACICIPRRRL--YVGKLEQWANKHNTETVSQGEAREGSEAP 1391
QY 62 -----DSKEDDLREAFIKSAAEFTLL-----ROYNSKNE----- 93
Db 1392 APEGKSSSGGKETPSDKLTAFIESAAVETFLMDRYKKEWLAOKRAELONGGLDYSS 1451
QY 94 -----DDKILHRMIPPEFRSMFYEGYRDICLDTDISKLAHDVTHAKKKITA 145
Db 1452 GQSDPDPNQKLLN-GVIPPDLRLMFYTLGDYRDILVH---GQNTSDSGNTGSMNNNT 1507
QY 146 VPOKIOSKTTNGK-----KVLREGE-----MWKEYGLSTWKGMCALSY 184
Db 1508 VLEASGNKEDMOKIOEKIOLIPKNGTLPVPKSSAQTPDKMNEHAESITWKMICALTY 1567
QY 185 -----NTEET-----KKMDE-----GVRTYLMKYYIKNDIKE--- 211
Db 1568 TEKNPPTISARGDENKTEKDEVEYEFKFGSTADKHGTAAPTGYTKQYDEKVKLEDTSG 1627
QY 212 -----YLEEFASRPPFLMWTYENGEDFVNKKRELVSLKRC----- 248
Db 1628 AKTPASASDPTPLSDVLRPPYFRYLEMGONGCKRKRRIKQIKEDMDGSKKYSGDG 1687
QY 249 DSGTLLNNGTSNKTCDNENCGACTQCEKRYKMMERMKHHSSQKKFQLYNSATY-- 306
Db 1688 EQCDRRD--TSNVSADLEG-RSCGNSCRFFYKWKIRKRREYDKQANAYS--KQTKYE 1742
QY 307 -----NNGLAVKEANSE 318
Db 1743 GSKGAGLNDHNKE 1755

RESULT 11

B71600
variant-specific surface protein 1 homolog PFBI055c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000

C:Accession: B71600
R:Gardner, M.J.; Tettejlin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.;
Petter, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; M0ID:99021743; PMID:9804551
A:Accession: B71600
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2197 <GAR>
A:Cross-references: GB:AE001434; GB:AE001362; NID:g38445341; PIDN:AAC71996.1; PID:g384
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFBI055c

Query Match 20.8%; Score 398.5; DB 2; Length 2197;
Best Local Similarity 25.4%; Pred. No. 1.8e-19;
Matches 133; Conservative 61; Mismatches 132; Indels 197; Gaps 21;

QY 15 KNNIHN-----CK-----KTEDAKMKCENT-----KLGEDGVCM 45
Db 918 KNAITNNDNLTDACKLKYGPGRKEREPPNKKCVSSGKSVATAGSSGATGKSGKAIQV 977
QY 46 PRRONLCVHYLTKL-----NDSKEDDLREAFIKSA 77
Db 978 PRRRRLYVGLTKLTSAAGTSSESGESSRASDVSGNGGDITTTESLRKWFJETAA 1037
QY 78 AETFL-----ROYNSKNE-----DD-----KILHRMIPPEFR 109
Db 1038 IETFLMHRKKEWAEOKRAELQRCNLLCTGASLNLGDDSNPOLOKSGTIPIDFLR 1097
QY 110 SMFYEGYRDICLDTDISKLAHDV-----TTAKKRTAVPQKI-----GSKIT 155
Db 1098 LMFYTLGDYRDILRGVADKNGGNNITLNASGNKDEKQMEKIOEKIOLPTSGKET 1157
QY 156 NG-----KVLREGEKWEYGLSTWKGMCALSY-----NTEK 189
Db 1158 RGPQNSVNDQSLMDLAELHVMHGMICALTYKDDNGLKGVAKPKQIENPEKIMETTK 1217
QY 190 K-MDEGVRYLYKYYIKNDIKE-----YLEEFASRPPFLMWTYENGEDV 234
Db 1218 KPKDE-----KYQYQAKLEDESGEKRPDSASGKTLDTFKRPYFRYLEMGENFC 1270
QY 235 KNRKKELVSLKRCDSCTLRNNG-----TSKTKDDNEN-----CGACKTQCEK 281
Db 1271 KRTEMLGKIKEDC-----YNGGRCSGDGLCKNEIYIDKIFGDLICTPCARHCRPYK 1326
QY 282 WM-----ERMKHHSSQKKFQLYKNSATYNNGL--AVKEANSE---TYKNDPEVTEA 329
Db 1327 WINTKRDENFKQSNAYSEQKKYEENDSAQKNNGVCGTLKDDAELNLRKNGPCKNES 1386
QY 330 NSAKHARDYL-----KTOLENNICTNG 351
Db 1387 EENKKADELDIFKRPDPTEKADNCRKCSSEFKICEHNHNCSSG 1429

RESULT 12

T14602
variant-specific surface protein - malaria parasite (Plasmodium falciparum) (fragment)
C:Species: Plasmodium falciparum
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jun-2000
C:Accession: T14602
R:Voss, T.S.; Felger, I.; Weiss, N.; Beck, H.P.
submitted to the EMBL Data Library, February 1998
A:Description: Identification of a conserved 5' flanking region of Plasmodium falciparum
A:Reference number: Z18158
A:Accession: T14602
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2135 <VOS>
A:Cross-references: EMBL:AF050740; NID:g2944094; PID:g2944095; PIDN:AAC05220.1
C:Genetics:
A:Gene: varph17

Query Match 20.8%; Score 397.5; DB 2; Length 2135;
 Best Local Similarity 27.5%; Pred. No. 2e-19;
 Matches 120; Conservative 51; Mismatches 149; Indels 117; Gaps 15;

10 SEQLEKNNHNCCK--TEDAKWCKCENTRLEDEG--VCMPPRONLCVHYLTAKLND-----61
 1407 SEGSEGGSHORAKRHTSDASGE--KSAKSGEPTGSGICIPRRRRKLVYGGITKAELQSS 1465
 62 -----DSKEEDLREAFIKSAAEFLLRQYNNKVNED-----94
 1466 QSOALSGOTTPTAGTPOADPLLAFAVESAAVETFFLMDRKKLNAPOSGSSLGAGAPLQ 1525
 95 -----DKILRDMIPPEFFRSMETFGDYDIDCT--DISERI-ADHDVTTA 139
 1526 LINGALISEETPEPISLKSGLTPDFLRLMYTLGDYDIDICVGVQDYIKALEAGDVKSS 1585
 140 KKKTAVFOKIGSKTTNG-----KVLREGMWKEYGLSIWKMICALSYNTE 187
 1586 KNPQOEISSKIEELIKNGTPTPTPTVTHSPSSGTTPTSSMWTNGOIHMKMICALTYKES 1645
 188 TKKMDGVR---TYLMKIY-----YKNNDIKF-----YLEEFASR 219
 1646 GEKKIEGVKATDNDLFEKLEKDKYSDYDKVYIEKENETEAKEGPDGLTPQTLLSNFVKR 1705
 220 PPLRLWTEGDEPFVKNRKEIVSLKKKC-----DSCTLRNNGTSNKTCDNE----267
 1706 PPFVRYLEEMGONCKTRKMLKDIIYECRNSDPNGDYGSGDGYHONNEIKHNNMHGI 1765
 268 NCAGCACTQCEKRYKMMERMKHVSQKKFOLYKNSATYNNGLAVKEANSEYKKNDEVT 327
 1766 DCRDCYKECKRKKYKWKVKKFDEFHNGKNKYGEEHEKLT-----NGDNVSGGDDNT 1815
 328 -----EANSAKHARDYK 340
 1816 NFOQIKKEKTAEDFLK 1832

RESULT 13

114029
 Variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
 N:Alternate names: erythrocyte membrane binding protein 1 (EMPI)
 C:Species: Plasmodium falciparum
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jun-2000
 C:Accession: T114029
 R:Chen, Q.; Barragan, A.; Fernandez, V.; Sundstrom, A.; Schlichtherle, M.; Sahlen, A.; C
 J. Exp. Med. 187, 15-23, 1998
 A:Title: Identification of Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1)
 A:Reference number: Z11860; MUID:98080592; PMID:9419207
 A:Accession: T114029
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-2228 <CHE>
 A:Cross-references: EMBL:AF003473; NID:q2961467; PID:q2961466; PIDN:AAOC05730.1
 C:Genetics:
 A>Note: FCR3S1.2-vari

Query Match 19.6%; Score 375; DB 2; Length 2228;
 Best Local Similarity 25.9%; Pred. No. 7.8e-18;

Matches 122; Conservative 46; Mismatches 139; Indels 164; Gaps 19;
 2 CKMVKLISEOIKNNHNCCKTEDA-----KMKC-----ENTKLGDEDC-VC 43
 886 CPYGVKVL-----KDN-----ESLQDACSLKYGGNNRKLGRKCYTPSGEPTTSSDKNCAIC 937
 44 MPPRRONLCVHYLTAKLND-----DSKEEDLREAFIKSA 78
 938 VPPRRRL---YIKKIYDAMKTESPOASGSEASTSGSTPPDSKA-LTKAFVESAAI 993
 79 EFTLLKQYNSK-----NVE-----DKILRDMIPPEFFRSMETFG 116
 994 ETEFLMHRKKEKKAVAOEGAGHGLPVEEGSPEDYDEK-LKEGKIPDGLKROMFYTIG 1052

117 DYRDICL-----DTDISEK-----IAHDVTTAKKTIYAV-----FOKIGS 152
 1053 DYRDILFSGSDNTTYSKTDTPSSNDNLKINVLIAASSTQDERKMKRYEIKNFRCS 1112
 153 KTNKKVLEREGWMEKYEGLSIWKMICALSYNETKKMDGYRTYLMK-----201
 1113 ERSAFNPVSHPOQWMEWENNGYIHWGVCAL---TSKRIKAKGYEKKRQKLENDENLMDEA 1169
 202 -----YIKNNDIKE-----YLEEFASRPPFLRWTEGDF 233
 1170 NKKRPPQYQYTWKVLIDENSGISPRYTQTAQSDNPTTLTHFVKRPTFRFMEEGESF 1229
 234 VKNRKEIVSLKKRCDSCTLRNNGTSNKTCDNE-----NCAGCACTQCE 277
 1230 CREKKRLKRIKDYCKA-----ENDDVGRCGDGACDSITHDYSTYSPNCGCGKHC 1285
 278 KYKMMERMKHVSQKKFOLYKNSATYNNGLAVKEANSEYKKNDEVT 328
 1286 SYRKWIERKKIEFHQGNAGQCKTDPATRNNGTDFKECKTLETPWDAK 1336

RESULT 14

128634
 Variant-specific surface protein 7 - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T28634
 R:Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Gulnet, F.; Herrfeldt, J.A.; Peterson,
 Cell 82, 89-100, 1995
 A:Title: The large diverse gene family var encodes proteins involved in cytoadherence
 A:Reference number: 220487; MUID:95330813; PMID:7606788
 A:Accession: T28634
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2182 <SUX>
 A:Cross-references: EMBL:LA2636; NID:g886379; PID:g886380; PIDN:AAA75399.1
 C:Genetics:
 A>Note: var-7

Query Match 19.6%; Score 374.5; DB 2; Length 2182;
 Best Local Similarity 25.4%; Pred. No. 8.2e-18;

Matches 121; Conservative 54; Mismatches 137; Indels 165; Gaps 19;
 1 PCKMVKLISEOIKNNHNC-----KKTEDAKMKC-----ENTKLGDEGVCMP 47
 851 PCQIVKLFED--DKSLKEACGLKYGEGEKPPMKCYTPSGVSTATSKDAICVPPR 908
 48 RQNLGVHYL-----TKLNDSEEDLREAFIKSAAEFLLRQYNS 89
 909 RRLRYGGLSQMASRGDETFTEVSESTAPSGSESEKLTARIESAIEFFLMHKYKE 968
 90 KNV-----ED--DKILRDMIPPEFFRSMETFGDYDIDCL--D 124
 969 EKKRPATQDAGLGLVSLPEPSPEGDPOTLOQTGVIPDFLRMYETLADYADIIYSGS 1028
 125 TDISERLADHDVTTAK-----KITAVFOKIGSKTTN 156
 1029 NDIISD-----TTGQTPSSNDNLKINYLEASGSTEQEKEMKQIQAKIKIILGANS 1081
 157 GKRVLEK-----EGWMEKYEGLSIWKMICALSY-----NTERKK--MDE 193
 1082 GVPPTVKNVSKYTPQYQWMEWENIADINMANVACALTYENADARGSATIEQKKDLKALMDE 1141
 194 GVR-TYLMKIY-----KNNDIKE--YLEEFASRPPFLRWTEGDFVKNRK 239
 1142 ANKNPTIEKQYTNVRLDESGAKSNDTIQPTLKNFVEIPTFRMLHEWGNFCFERAK 1201
 240 ELVSLKKKDCSCLRNNGTSNKTCDNEN-----GACTQCEKRYKMM 284
 1202 RLAKIHKEC-----MDGEKQYSGDGEYCEEFISKQYVNLQDLSSCAKPCPLYTWIE 1256
 285 RMKRRHVSQKKFOLYKNSATYNNGLAVKEANSEYKKNDEVTANSAKHARDYKLT 341

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein protein search, using sw model

Run on: June 20, 2003, 14:37:29 ; Search time 15.4461 Seconds

(Without alignments)
942.514 Million cell updates/sec

Title: US-10-087-013-9

Perfect score: 1913
Sequence: 1 PCRMVOKLISEQIEKNHIN.....AKHARDLKTLENMICTNG 351

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|---------------------|
| 1 | 329 | 17.2 | 1435 | EBAL_PLAFC | P19214 plasmodium |
| 2 | 269 | 14.1 | 1070 | PVDR_PLAVS | P22230 plasmodium |
| 3 | 261 | 13.6 | 1153 | PVDB_PLAKN | P20493 plasmodium |
| 4 | 231 | 12.1 | 1073 | PVDA_PLAKN | P22545 plasmodium |
| 5 | 225 | 11.8 | 1070 | PVDG_PLAKN | P50494 plasmodium |
| 6 | 122.5 | 6.4 | 607 | HTPG_PLAKN | O8r0h4 fuscobacteri |
| 7 | 117 | 6.1 | 1391 | MST2_DROHY | O86866 drosophilla |
| 8 | 115 | 6.0 | 944 | NUP1_YEAST | P21380 saccharomyc |
| 9 | 114 | 6.0 | 993 | SCP1_MOUSE | O62209 mus musculu |
| 10 | 110 | 5.8 | 1995 | YCXY_CHIRE | P36495 chlamydomon |
| 11 | 109 | 5.7 | 845 | SCP1_MESAU | O60563 mesocricetu |
| 12 | 106.5 | 5.6 | 896 | RA50_THEAC | O97050 thermoplasma |
| 13 | 106 | 5.5 | 767 | TOP1_CRIGR | O07050 cricetus |
| 14 | 105.5 | 5.5 | 2869 | RBP1_PLAVB | O00798 plasmodium |
| 15 | 104.5 | 5.5 | 817 | ENPL_CARRO | P35016 catharanthu |
| 16 | 103.5 | 5.4 | 2230 | GOG4_HUMAN | Q13439 homo sapien |
| 17 | 102.5 | 5.4 | 1976 | MYHA_BOVIN | Q27991 bos taurus |
| 18 | 101.5 | 5.3 | 845 | VAV_HUMAN | P15498 homo sapien |
| 19 | 101.5 | 5.3 | 997 | SCP1_RAT | Q03410 rattus norv |
| 20 | 101.5 | 5.3 | 1160 | TFC3_YEAST | P34111 saccharomyc |
| 21 | 101.5 | 5.3 | 1727 | ALMT_SCHPO | Q0475 schizosacch |
| 22 | 101.5 | 5.3 | 1976 | MYHA_HUMAN | P35580 homo sapien |
| 23 | 101.5 | 5.3 | 2339 | RPCL_PLAFA | P27625 plasmodium |
| 24 | 101 | 5.3 | 1972 | MYHB_HUMAN | P37749 homo sapien |
| 25 | 100.5 | 5.3 | 581 | Y208_BORBU | O51226 borrelia bu |
| 26 | 100.5 | 5.3 | 1928 | MYSL_YEAST | P08964 saccharomyc |
| 27 | 100 | 5.2 | 634 | NOG1_MOUSE | Q09969 mus musculu |
| 28 | 99.5 | 5.2 | 843 | VAV_RAT | P41100 rattus norv |
| 29 | 99 | 5.2 | 671 | RPSP_HELIPY | P55993 helicobacte |
| 30 | 98.5 | 5.1 | 1790 | USOL_YEAST | P25386 saccharomyc |
| 31 | 98 | 5.1 | 443 | CSN2_HUMAN | O15647 homo sapien |
| 32 | 98 | 5.1 | 1539 | Y373_HUMAN | O15078 homo sapien |
| 33 | 98 | 5.1 | 1682 | MSP1_PLAFC | P19598 plasmodium |

| | | | | | | |
|----|------|-----|------|---|-------------|--------------------|
| 34 | 98 | 5.1 | 1805 | 1 | HMW2_MYCGE | P47460 mycoplasma |
| 35 | 98 | 5.1 | 2136 | 1 | YC2F2_MARPO | P09975 methanilla |
| 36 | 97.5 | 5.1 | 441 | 1 | Y072_METUA | O60306 methanococ |
| 37 | 97.5 | 5.1 | 519 | 1 | CPT7_RANDOM | O57525 rana dybows |
| 38 | 97.5 | 5.1 | 726 | 1 | HS9A_BRARE | O90474 brachydanio |
| 39 | 97 | 5.1 | 609 | 1 | YSW1_YEAST | P38280 saccharomyc |
| 40 | 97 | 5.1 | 807 | 1 | HYSA_STRAU | O59801 staphylococ |
| 41 | 97 | 5.1 | 976 | 1 | SCP1_HUMAN | O15431 homo sapien |
| 42 | 97 | 5.1 | 1972 | 1 | MYHB_MOUSE | O08638 mus musculu |
| 43 | 96.5 | 5.0 | 795 | 1 | ENPL_CHICK | P08110 gallus gall |
| 44 | 96.5 | 5.0 | 845 | 1 | VAV_MOUSE | P27870 mus musculu |
| 45 | 96.5 | 5.0 | 1005 | 1 | RA50_METUA | O58718 methanococ |

ALIGNMENTS

| RESULT 1 | EBAL_PLAFC | STANDARD | PRT | 1435 AA |
|-----------------------|--|----------|-----|---------|
| AC | P19214 | | | |
| DT | 01-NOV-1990 (Rel. 16, Created) | | | |
| DT | 01-FEB-1996 (Rel. 33, Last sequence update) | | | |
| DT | 01-FEB-1996 (Rel. 33, Last annotation update) | | | |
| DE | Erythrocyte-binding antigen EBA-175. | | | |
| OS | Plasmodium falciplarium (isolate Camp / Malaysia). | | | |
| OC | Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium. | | | |
| OX | NCBI_Taxid=5835; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=90377299; PubMed=2204835; | | | |
| RA | Sam B.K.L.; | | | |
| RT | "Sequence conservation of a functional domain of erythrocyte binding antigen 175 in Plasmodium falciplarium." | | | |
| RL | Mol. Biochem. Parasitol. 41:293-296(1990). | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch). | | | |
| DR | EMBL: X52524; CAA36756.1; - | | | |
| KW | PIR: S11561; S11561. | | | |
| FT | DOMAIN 159 1104 | | | |
| FT | ESSENTIAL FOR BINDING TO ERYTHROCYTES. | | | |
| FT | VARIANT 1031 1031 E -> V (IN STRAINS FCR-3 AND ITG). | | | |
| SO | SEQUENCE 1435 AA; 167389 MW; 32A4309021B1C3D6 CRC64; | | | |
| Query Match | 17.28; Score 329; DB 1; Length 1435; | | | |
| Best Local Similarity | 24.88; Pred. No. 2.9e-15; | | | |
| Matches | 87; Conservative 63; Mismatches 113; Indels 88; Gaps 13; | | | |
| QY | 14 EKNNHCKKTEDEAKKCKEYKGEDEGVCMPPRRORLCAHYLTKLNDDS-----KEEDLR 69 | | | |
| DB | 463 DKNSV-----DTNKKWPECKNPYIISTDVCPVPRROELCGNIDRLTKMLIKETILA 518 | | | |
| QY | 70 EAFKSAAEFTFLROYNKRKNVEDKILRDMIPREFRSMFTFGDYRDICLDPISE 129 | | | |
| DB | 519 IAIYES-----RLKRRYKKNDDKDKVCIINK-----TFADIDIDIGGTQYWN 561 | | | |
| QY | 130 KIADHDVYTKAKKITAVFOKIGSTTKGKV-----LEREGWKEYGLSTWKGMLCA 181 | | | |
| DB | 562 DLSNRKL-----VGKINTSKKYVHRKRNKDKLEFRDWMVVIKKDVNN----- 603 | | | |
| QY | 182 LSVTEFKMKDEGVRTLYMKYIYKNNDI-KEYLEEFASRPPEFLRWVTEWGEDFVKNRKE 240 | | | |
| DB | 604 -----VISVFEKDKVYCE--DDIENIQOFFFWSEKGDVYQDKTKM 644 | | | |
| QY | 241 LVSLKKKDCSTLRNNGTSNKTCDNENGCACKTQCEKRYKWRMRKHYSSOKKKFQLY 300 | | | |

Db 645 IETLKEC-----KEKPEDN-----CKSKNSYKEWISKKEEYNQAKQYEX 690
 QY 301 KNSATNNGLAWEANSEY--KNDEVEANSAKARLYLTQLENMCT 349
 Db 691 QKONNKKMSEFSISPEYLLKYSEKCSNLNFEDEKEELHSDYKNK-CT 740

RESULT 2

PVDR_PLAVS
 ID PVDR_PLAVS STANDARD: PRT: 1070 AA.
 AC P22290:
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Duffy receptor precursor (Erythrocyte binding protein).
 GN PVDR.
 OS Plasmodium vivax (strain Salvador I).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=126793;

SEQUENCE FROM N.A.

MEDLINE=91187056; PubMed=1849231;
 Fang X., Kaslow D.C., Adams J.H., Miller L.H.;
 RT "Cloning of the Plasmodium vivax Duffy receptor";
 RL Mol. Biochem. Parasitol. 44:125-132(1991).
 CC - FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP
 CC DETERMINANT.
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - SIMILARITY: HIGH, TO P. KNOWLESI DUFFY RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC or send an email to license@lsb-sib.ch).

EMBL: M61095; AAA63423.1;

DR Malaria; Receptor; Glycoprotein; Signal; Transmembrane.
 KW SIGNAL 1 20
 FT CHAIN 21 1070
 FT DOMAIN 21 1007
 FT TRANSMEM 1008 1025
 FT DOMAIN 1026 1070
 FT CARBOHYD 183 183
 FT CARBOHYD 255 255
 FT CARBOHYD 351 351
 FT CARBOHYD 420 420
 FT CARBOHYD 715 715
 FT CARBOHYD 787 787
 FT CARBOHYD 825 825
 FT CARBOHYD 903 903
 FT CARBOHYD 938 938
 SQ SEQUENCE 1070 AA; 119683 MW; CB051DF13E294603 CRC64;

Query Match 14.13; Score 269; DB 1; Length 1070;
 Best Local Similarity 26.28; Pred. No. 3e-11;

Matches 84; Conservative 50; Mismatches 111; Indels 76; Gaps 15;

QY 6 QKLSIOI-----EKNNHNC--KTEADAKKMCENTKLGDEGVCMPPRONLCVHYL 56
 Db 196 KTLISSAIIINHAFLQNTVMKNCNKKRRERDMDC-NFK--KDVCIPIRRYQLCKWEL 250
 QY 57 TKL--NDSD--KEEDIREAFIK-----SAAAEFTLLROYNSKNVEDDKILHRDMPPE 106
 Db 251 TNLVNNNDTDFNHFHDIKRLKLYDAVAGDLLLKLNTRYNK----- 298

QY 107 FFSMFTFGDYGDIQDLDIS-----EKIDHVTYAKKKTITVAFKISGTTNGKVVLE 162
 Db 299 FCDIDIRKSLDGFQIDTGMETIGYSKYVENN-----LRSIF-----GTDERAQR 345

QY 163 REGWKEVGLSIVKMCALSNTEKMKDEGVRTYLLKYYTKNNDIKYLEEPASRPF 222
 Db 346 KRWOMEKRAQIWTIMMYSV-----KKRLGNFWICKLVAVN-----IPEOI 389
 QY 223 LRMVTEWGEDFYKNNKKEELYSLKJBNNGTSKTKTDDDNENGACACKTOCEKRYKW 282
 Db 390 YRMIRKMRDYSSELPTEVOKLEKCDG--KINTDKVC-----KVPCCNACKSYDOW 442

QY 283 MERWKKHYSQKKRFRQLYKNS 303
 Db 443 ITRKKNQDVLNKNKFIYSKNA 463

RESULT 3

PVDR_PLAVS
 ID PVDR_PLAVS STANDARD: PRT: 1153 AA.
 AC P50493:
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Duffy receptor, beta form precursor (Erythrocyte binding protein).
 OS Plasmodium knowlesi.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5850;

SEQUENCE FROM N.A.

MEDLINE=92357776; PubMed=1496004;
 RA Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;
 RT "A family of erythrocyte binding proteins of malaria parasites";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089(1992).
 CC - FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP
 CC DETERMINANT.
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - SIMILARITY: HIGH, TO P. VIVAX DUFFY RECEPTOR.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC or send an email to license@lsb-sib.ch).

EMBL: M90694; AAA29603.1;

DR Malaria; Receptor; Glycoprotein; Signal; Transmembrane;
 KW Multigene family.
 FT CHAIN 1 21
 FT SIGNAL 22 1153
 FT DOMAIN 22 1085
 FT TRANSMEM 1086 1106
 FT DOMAIN 1107 1153
 FT CARBOHYD 134 134
 FT CARBOHYD 179 179
 FT CARBOHYD 202 202
 FT CARBOHYD 252 252
 FT CARBOHYD 348 348
 FT CARBOHYD 430 430
 FT CARBOHYD 467 467
 FT CARBOHYD 576 576
 FT CARBOHYD 626 626
 FT CARBOHYD 722 722
 FT CARBOHYD 847 847
 FT CARBOHYD 856 856
 FT CARBOHYD 900 900
 FT CARBOHYD 910 910
 FT CARBOHYD 935 935
 SQ SEQUENCE 1153 AA; 130471 MW; 6497BD16CE7BEC01 CRC64;

Query Match 13.6%; Score 261; DB 1; Length 1153;
 Best Local Similarity 23.9%; Pred. No. 1.2e-10;
 Matches 87; Conservative 50; Mismatches 125; Indels 102; Gaps 14;

QY 3 KMWOKLSIQ--IEKNNHNC--KTEADAKKMCENTKLGDEGVCMPPRONLCVHYLTK 58

```

Db      195  RMCSGVINOTFLQKNVRRCKRRKGTGTRDWCPTK-----DVCIPDRYQLCMKELTN 249
      59  LNDSSKEE-----DLREAFIKSAAEFFLLQYNSKNVVEDDKILHRMIPPEFF 108
      250  LVNNTKTHSHNDITFLKLNKEKLYDAVEGDLTKKNYNSSED-----LC 297
      109  RSMFYEGDYRDICLDTDISSEKADHDVYTAKKITAVFOKIGSKTKNGKRYLEREGWKK 168
      298  KDIKMSLEDPGDIIMGTDM-EGIGYSQV--VENNLTFV-----GTGKTQIDRRKKWN 348
      169  EYGLSIWGMALCALSYNTETKKMDGVRYYLKKYIKNNDIKYLEEFASRPPLRWYTE 228
      349  ESKKYIWEATILISV-----KKKNGYSAMNCK-----EDVQINVE-----POIYRWIRE 392
      229  WGEDFYNNKKKELYSIKKKCD-----CTLRNNGTSNKTCDNENCGACKTOCEKY 279
      393  WGMADYSELPDEKRIKKECDKRLYTNLRICTM-----SPCNDCKLY 436
      280  KKMERRKKHYSSQKKKPOLYKNSATYNNGLAVKEANSEYTKNDPEVTEANSKHAARDYL 339
      437  DQWITKKKQWDLSTKFFSSVKKG-----QIETENTTTAYDIL 475
      340  KYOL 343
      476  KOEL 479
      479  KOEL 479

RESULT 4
PVDA_PLAKN STANDARD; PRT; 1073 AA.
AC P22545;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Duffy receptor, alpha form precursor (Erythrocyte binding protein).
OS Plasmodium knowlesi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5850;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92357776; PubMed=1496004;
RA Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;
RT "A family of erythrocyte binding proteins of malaria parasites.";
RN Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089(1992).
RP SEQUENCE OF 296-1073 FROM N.A.
RX MEDLINE=91004213; PubMed=2170017;
RA Adams J.H., Hudson D.E., Torii M., Ward G.E., Wellens T.E.,
RT Alkawa M., Miller L.H.;
RT "The Duffy receptor family of Plasmodium knowlesi is located within
RT the micronemes of invasive malaria merozoites.";
RN Cell 63:141-153(1990).
CC -|- FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP
CC DETERMINANT.
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
CC -|- SIMILARITY: HIGH, TO P.VIVAX DUFFY RECEPTOR.
CC -|- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M90466; AAA29602.1; -
CC EMBL: M68517; AAA29590.1; -
CC EMBL: M68518; AAA29591.1; -
CC PIR: A35970; A35970.
CC Malaria; Receptor; Glycoprotein; Signal; Transmembrane;
CC Multigene family.
CC SIGNAL 1 21 POTENTIAL.

```

```

FT CHAIN 22 1073 DUFFY RECEPTOR, ALPHA FORM.
FT DOMAIN 22 1007 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1008 1029 POTENTIAL.
FT DOMAIN 1030 1073 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 348 348 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 746 746 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 779 779 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 788 788 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1073 AA; 120683 MW; 3965FC9F46B71808 CRC64;

Query Match 12.1%; Score 231; DB 1; Length 1073;
Best Local Similarity 22.8%; Pred. No. 1,3e-08;
Matches 86; Conservative 59; Mismatches 146; Indels 86; Gaps 15;

      3  KMYQKLISEQIEKNNI-----HNCKRTEDAKKCCENTKLGDEGVCMPPRRONICVHYLTK 58
      195  RMCSGVINOTFLQKNVRRCKRRKGTGTRDWCPTK-----AEKICISVRRYQLCMGELTN 249
      59  LNDSSKEEDIRE-APIKSAAEFFLLQYNSKNVVEDDKILHRD--MIPPEFRSMFYTF 115
      250  LVNNTKTHSHNDITFLK-----LNLKRKIMYDAVEGDLTKKNYNNYNEFCEDIRWGL 304
      116  GDYRDICLDTDISSEKADHDVYTAKKITAVPQIKSGTKNGKRYLEREGWKKYGLSIW 175
      305  GDFGDIIMGTNM-EGIGYSQV--VENNLROVF-----GTDEKAKODRKQKNWSEKHEHW 355
      176  KGMICALSYNTETKKMDGVRYYLKKYIKNNDIKYLEEFASRPPLRWYTEMGEDFYK 235
      356  RAMMFSIRSLRKEK-----FWICK-KVYTLKVE-----POIYRIRRMGNDYMS 399
      236  NRKKEVSLKKKDCSCTLRNNGTSNKTCDNENCGACKTOCEKYKKMERRKKHYSSQKK 295
      400  KLPKEQGLKNEKCAKSLYYNN--MAIC-----MLPCHDACKSDYITKKKQWVLYST 452
      296  KFOLYKNS-----ATYNNGL-----AVKEANSE 318
      453  KFSVKKTKQKIGTENIATAYDILKQELNGKKEATFENEIKRKNLYNHLCPVVEARKN 512
      319  TYKNDPEVTEANSKAKHA 335
      513  TQENKVNNGSGVESKAA 529

RESULT 5
PVDA_PLAKN STANDARD; PRT; 1070 AA.
AC P50494;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Duffy receptor, gamma form precursor (Erythrocyte binding protein).
OS Plasmodium knowlesi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5850;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92357776; PubMed=1496004;
RA Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;
RT "A family of erythrocyte binding proteins of malaria parasites.";
RN Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089(1992).
CC -|- FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP
CC DETERMINANT.
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
CC -|- SIMILARITY: HIGH, TO P.VIVAX DUFFY RECEPTOR.
CC -|- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M90695; AAA29604.1; -
KW Malattia; Receptor; Glycoprotein; Signal; Transmembrane;
MW
KW Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 1070
FT TRANSMEM 22 1003
FT TRANSMEM 1004 1025
FT DOMAIN 1026 1070
FT CARBOHYD 134 134
FT CARBOHYD 179 179
FT CARBOHYD 676 676
FT CARBOHYD 743 743
FT CARBOHYD 785 785
FT CARBOHYD 936 936
FT SITE 279 281
SEQUENCE 1070 AA; 120931 MW; 703D6881BC11B50 CRC64;

POTENTIAL.
DUFFY RECEPTOR, GAMMA FORM.
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
CELL ATTACHMENT SITE (POTENTIAL).
Query Match 11.8%; Score 225; DB 1; Length 1070;
Best Local Similarity 20.3%; Pred. No. 3.3e-08;
Matches 80; Conservative 53; Mismatches 135; Indels 126; Gaps 13;

QY 3 KAVOKLISEQIEKNTI---HNCKTEDAKKCENTKLGDEGCMPPRONLCVHLTK 58
DB 195 KMSGGAINGIFLQNNVNDKCDKRRKGEWDCTEK---DVCIDRRYQLCMEITN 249
QY 59 LNDSDS-----KEEDLEAFIKSAAEFTFLROYNSKNVDDDLHRDMIPPEFR 109
DB 250 LVDTDTHRHSDIIRKSKSRRLIYDVGGRDILLKRYNNVSD-----LCK 297
QY 110 SMFTFGDYRDCIDLTIDISEKIDADHYTAKKITYAPQKIGSKTNGKYLEREGWKE 169
DB 298 DIKSLDQDFDIIINGTDM-EGIGSLV--VONNLRSIF---CTGTSALDRKKWMD 348
QY 170 YGLSIWKGMLCAL----YNTETKKMGEGVRYTIKTYKNNDIKEVLEEFASPPFLRW 225
DB 349 HKKDIWKAMILSVKERNYSAMNCKEDVQIKV-----EPQIYRW 387
QY 226 VTEGDEFEVKNRKEKELYSKKKCD----CTLRNNGTSNKTGDDNENGCACKTQC 276
DB 388 IRENGRDYSEFRDQRRLNKEKCDKLYSTMLICTL-----PPCNMAC 431
QY 277 EYKKMMERKKKHHSSQKKFQYK-----NSATYNNGL- 310
DB 432 KSYDEMTIGKKRQMDVLTSTKFSYVKKAKIETENIARAYDILKQELNGFNEVTEENIK 491
DB 311 -----AVKEANSEYKNDPEVTEANSK 333
DB 492 RDKLYNYTCVCIQDEAKNKTOENKKNVSGVEAK 525

RESULT 6

HTPG_FUSNN STANDARD: PRT: 607 AA.
AC Q08GH4 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chaperone protein htpg (Heat shock protein htpg) (High temperature protein G)
DE HTPG OR FN0321.
OS Fusbacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusbacterium; Fusbacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N. A.
RC STRAIN-ATCC 25586;
RX MEDLINE=2186934; PubMed=11889109;
RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,

RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vastisa O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,
RA Larsten N., D'Souza M., Malinas T., Pusch G., Haselkorn R.,
RA Founsein M., Kyrides N., Overbeek R.,
RT "Genome sequence and analysis of the oral bacterium Fusbacterium
RT nucleatum strain ATCC 25586."
RT J. Bacteriol. 184:2005-2018(2002).
CC - FUNCTION: Molecular chaperone. Has ATPase activity (By
CC similarity).
CC - SUBUNIT: Homodimer (By similarity).
CC - SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC - SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AEO10544; AA194527.1; -
DR PROSITE: PS00298; HSP90; 1.
KW Chaperone; ATP-binding; Heat shock; Complete proteome.
FT DOMAIN 1 323 A: SUBSTRATE-BINDING (BY SIMILARITY).
FT DOMAIN 324 534 B (BY SIMILARITY).
FT DOMAIN 535 607 C.
SQ SEQUENCE 607 AA; 70368 MW; D1519D82B0FD1B08 CRC64;

Query Match 6.4%; Score 122.5; DB 1; Length 607;
Best Local Similarity 24.1%; Pred. No. 0.21;
Matches 66; Conservative 48; Mismatches 121; Indels 39; Gaps 11;

QY 101 DMIPPEFRSMFTFG---DYRDIQDIDDISKINDHYTT--AKKITYAPQKIGSKT 154
DB 299 DLI-PEYFN--FTSGVDDSLNLSRIIDQONAELOYISKNEKTTSEKLT--LK 352
QY 155 TNGKYLREKGMKREYGLSIWKGMLCALSYNETKMD-----EGVRYLMKYKN 206
DB 353 NDREKYE--FWKEGRSIRKAGVDMFGNKR-KIQLDILLFVSHDDKYTLKEVDMM 408
QY 207 NDIKEYLEEFASPPFLRWYTEGDEFEVKNRKEKELYSKKKCDSCILRN----- 255
DB 409 GDNKEILYPAESVDKARYLPKM--EKLEQGEVILLIDKIDFTLMAMRDYSGKEFS 466
QY 256 -NGTSNKTGDDNENGCACKTQCEKRYKMKMRKHHSSQKKFQYKNSATYNNGLAVE 314
DB 467 INSDKFEPSDKKEKEEYKRIADENKELIEKAKEFLDKYSEVLSNNIGNSASSLLANG 526
QY 315 ANS-ETYYKNDPEVTEANSKHAHDYLTQLENNI 347
DB 527 GLSLEMEKTLSEMTNNNDMPKAEKVALAINPEHVL 560

RESULT 7

MST2_DROHY STANDARD: PRT: 1391 AA.
AC Q08696;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Axoneme-associated protein mst101(2).
GN Mst101(2).
OS Drosophila hydei (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7224;
RN [1]
RP SEQUENCE FROM N. A., AND CHARACTERIZATION.
RC MEDLINE=9504538; PubMed=7957199;
RX Neesen J., Padmanabhan S., Buenemann H.,
RA "Randomly arranged repeats of a novel highly charged 16-amino-acid


```

Db 257 SDVQSLKLSKDEDELKNTL-INNELNELKNAEERKQLEFEKKNELRKRTNELNLKIKSD 315
OY 78 AETFLROYNSKNVEDDKILHRDMLPPEFFRSMFTFDYDIDLDITDISE-----KTA 132
Db 316 MDQLOKQONKESRIKDE--LNE-----LTFPSNGSSOSAK 351
OY 133 DHDVTTAKKITAFAVOKISGTTNGKVLREGWMEYGLSIWGMICALNTERTKMD 192
Db 352 ENELMKLNKRIAELEEEISTKNS--QLAKEG-----KLASLMQLOLEKLNROSDQ 403
OY 193 EGVRYFLMKYIKNN-----DIKEYIEEASAPPELWVTENGEDFVKNRKKEVLSLKK 247
Db 404 LGSREBELK--KTNDKLOKODIRIAREEIVS-----KDERIIDLOK 442
OY 248 C-----DSCTLRNNGTSNKTCDN-----ENCGACKTOCEKTKKMERBK-KH 289
Db 443 VKOLENDLFIKKTHSEKSTIINDNESEKDLIKILLEN--DLKVAQEKYSKKEKELEKE 500
OY 290 YSSQKKRPOLYKNSATYNNGLAVKEANSETYNDEPTEANSKARDYLKTOLENM 346
501 FVYKISEKLEDEKTTLMNEKISNLAEANSOLKNKIE-DNSTATHMKREYERQLES 556

```

RESULT 9

SCPL_MOUSE

ID SCPL_MOUSE STANDARD: PRT: 993 AA.

AC 062209; 062329; 009205; P70192;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Synaptonemal complex protein 1 (SCP-1 protein).

GN SCPL OR SCPL

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.

NCBI_TaxID=10090;

[1]

SEQUENCE FROM N.A.

RC STRAIN-CBA; TISSUE-Testis;

RX MEDLINE=96004899; PubMed=7548215;

RA Sage J., Martin L., Cuzin F., Rassoulzadegan M.;

RT "cDNA sequence of the murine synaptonemal complex protein 1 (SCP1).";

RL Biochim. Biophys. Acta 1263:258-260(1995).

[2]

SEQUENCE FROM N.A.

RC STRAIN-Swiss; TISSUE-Testis;

RA Kerr S.M., Taggart M.H., Lee M., Cooke H.J.;

RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.

[3]

SEQUENCE OF 1-149 FROM N.A.

RA STRAIN-C57BL/6; Sage J., Li Y., Martin L., Mattei M.-G., Guenet J.-L., Liu J.G.,

RL Hoog C., Cuzin F., Rassoulzadegan M.;

[4]

Submitted (Jan-1997) to the EMBL/GenBank/DBJ databases.

[5]

SEQUENCE OF 95-787 FROM N.A.

RA STRAIN-ICR; TISSUE-Testis;

RL Tsuchida J., Nishida Y., Nozaki M., Uchida K., Nishimune Y.;

[6]

Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

[7]

FUNCTION: MAJOR COMPONENT OF THE TRANSVERSE FILAMENTS OF

SYNAPTONEMAL COMPLEXES (SCS), FORMED BETWEEN HOMOLOGOUS

CHROMOSOMES DURING MEIOTIC PROPHASE.

[8]

SUBCELLULAR LOCATION: NUCLEAR. IN TRIPARTITE SEGMENTS OF

SYNAPTONEMAL COMPLEXES, BETWEEN LATERAL ELEMENTS IN THE NUCLEUS.

[9]

FOUND ONLY WHERE THE CHROMOSOME CORES ARE SYNAPSED. ITS N-TERMINUS

C-TERMINUS EXTENDS WELL INTO THE LATERAL DOMAIN OF THE

SYNAPTONEMAL COMPLEX (BY SIMILARITY).

[10]

DOMAIN: CONSISTS OF AN ALPHA-HELICAL STRETCH OF 700 AA RESIDUES,

FLANKED BY N- AND C-TERMINAL GLOBULAR DOMAINS. THE C-TERMINAL

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[11]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[12]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[13]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[14]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[15]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[16]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[17]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[18]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[19]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[20]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[21]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[22]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[23]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[24]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[25]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[26]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[27]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[28]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[29]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[30]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[31]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[32]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[33]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[34]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[35]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[36]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[37]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[38]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[39]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[40]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[41]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[42]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[43]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[44]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[45]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[46]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[47]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[48]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[49]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[50]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[51]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[52]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[53]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[54]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[55]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[56]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[57]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[58]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[59]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[60]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[61]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[62]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[63]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[64]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[65]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[66]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[67]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[68]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[69]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[70]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[71]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[72]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[73]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[74]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[75]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[76]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[77]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[78]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[79]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[80]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[81]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[82]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[83]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[84]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[85]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[86]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[87]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[88]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[89]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[90]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[91]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[92]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[93]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[94]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[95]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[96]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[97]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[98]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[99]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

[100]

DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

```

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: Z38118; CAAB6262.1; ALT. INIT.
DR EMBL: L41069; AAA64514.1; ALT. INIT.
DR EMBL: U62864; AAC53335.1; JOINED.
DR EMBL: U62864; AAC53335.1; JOINED.
DR EMBL: U62861; AAC53335.1; JOINED.
DR EMBL: U62862; AAC53335.1; JOINED.
DR EMBL: U62863; AAC53335.1; JOINED.
DR EMBL: D86539; BAA13639.1; JOINED.
DR MGD: MGI:105931; Scp1.
KW Nuclear protein; Meiosis; Cell division; Phosphorylation;
KW DNA-binding; Coiled coil.
FT DOMAIN 12 97 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 104 815 COILED COIL (POTENTIAL).
FT DOMAIN 114 117 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 697 700 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 898 901 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 978 986 ARG/LYS-RICH (BASIC).
FT CONFLICT 527 527 F -> L (IN REF. 2).
SQ SEQUENCE 993 AA; 115962 MW; 1A4FA790D6FAFE6 CRC64;

```

Query Match Best Local Similarity 6.0%; Score 114; DB 1; Length 993.

Matches 82; Conservative 63; Mismatches 163; Indels 112; Gaps 15;

```

OY 2 CKWYQXLSIPQIEKNH-----NCKKT-----EDAKKCEMT-----K 35
Db 180 CNLIKETCAESAERTNKEYEERETROYVDLSNIEKMLAEELRVQAEANRLWEHFK 239
OY 36 LGEDEGVCMPRRONLCVHYLTFLNDSKEEDIREFAFKSAAE-----TFLLROYNS 89
Db 240 LKEDH-----EKIQHEEYQKEVN--NKNQVSELLQSAEENKKDLTFLEESRDK 292
OY 90 KNVEDDKILHRDMLPPEFFRSMFTFDYDIDLDITDIS--EKIADHDVTTAKKITT-- 144
Db 293 ANOLEEKTQLQDENLKLSEKDKHITSELEDIMQSMQSMSTQKALEEDILQIATKTISI 352
OY 145 -----AVFQKIG--SKTNGKVKVLEREGWMEYGLSIWGMICALS--YNTETPKMDEG-- 194
Db 353 TEYKEMQELNNAKTHSVTFE-----LKATTCYLEELRTEQORLEKNEB 400
OY 195 -VRYTLMKYIYKNDIKLEYLEFPASRPPLRWVTENGEDFVKNRKKEVLSLK----- 245
Db 401 QLKRLITVELOKRSNLEEMTK-----FKNNKEVELEIKNILLAEODK 442
OY 246 -----KKDCSCTLRNNGTSNKTCDNENCGACKTOCEYKTKMERBK 289
Db 443 LDEKROVEKLAELQEKDELTFLLTREKEVHDLQEOVYTVKTSQHYLKOVEREMKTE 502
OY 290 YSSQKKRPOLYKNSA--TYNGLAVKEANSETYNDEPTEANSKARDYLKTOLENM 346
Db 503 LEKELKLNELTASCOMLLEKKKFFQEAASDMLAEIKKNOEDILNCKQKEEFLNQIETNL 562

```

RESULT 10

YCX7_CHLRE

ID YCX7_CHLRE STANDARD: PRT: 1995 AA.

AC P36495; Q37303;

DT 01-JUN-1994 (Rel. 29, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 232.2 kDa protein (ORF1995) (ORF-S) (ORFA).

OS Chlamydomonas reinhardtii.

OG Chloroplast.

OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

NCBI_TaxID=3055;

This SWISS-PROT entry is copyright. It is produced through a collaboration

```

RP [1] MEDLINE FROM N.A.
RX MEDLINE-97218038; PubMed=9065699;
RA Boudreau E., Turmel M., Goldschmidt-Clermont M., Rochemaix J.-D.,
RA Sivan S., Michaels A., Lelu S.;
RT "A large open reading frame (orf1995) in the chloroplast DNA of
RT Chlamydomonas reinhardtii encodes an essential protein.";
RL Mol. Gen. Genet. 253:649-653(1997).
RN [2]
RP SEQUENCE OF 1573-1995 FROM N.A.
RC STRAIN-CW15;
RA Sivan S., Michaels A.;
RN Submitted (DEC-1993) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 1925-1995 FROM N.A.
RX MEDLINE-87031585; PubMed=2876928;
RA Moessner J.P., Gilham N.W., Boynton J.E.;
RT "The sequence of the chloroplast atpB gene and its flanking regions
RT in Chlamydomonas reinhardtii.";
RL Gene 44:17-28(1986).
RN [4]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE-94207185; PubMed=8155879;
RA Boudreau E., Otis C., Turmel M.;
RT "Conserved gene clusters in the highly rearranged chloroplast genomes
RL of Chlamydomonas moewusii and Chlamydomonas reinhardtii.";
RL Plant Mol. Biol. 24:585-602(1994).
CC -I- FUNCTION: ESSENTIAL FOR CELL GROWTH. MAY BE INVOLVED IN BINDING
CC CHLOROPLAST DNA TO EITHER THE CHLOROPLAST ENVELOPE OR THE
CC THYLAKOID MEMBRANE.
CC CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC CC -I- SIMILARITY: TO C.VULGARIS ORF819.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR DR EMBL: X92726; CAA63385.1; -
DR EMBL: X76934; CAA54257.1; -
DR PIR: S41289; S41289.
DR InterPro: IPR001123; IySs.
DR Chloplast; Transmembrane; Thylakoid.
FT FT TRANSMEM 31 51 POTENTIAL.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 212 232 POTENTIAL.
FT TRANSMEM 254 274 POTENTIAL.
FT TRANSMEM 307 327 POTENTIAL.
FT CONFLICT 1588 1588 K->R (IN REF. 2).
SQ SEQUENCE 1995 AA; 232194 MW; E53AF167EAF5F08A CRC64;

```

```

Db      1214 KSSLVYNOTQLLRVATDDLDQWRTKQVITRKRSARRDRPKRKQIAVYNNKLALSLSKVE 1273
Oy      188 TKKM-----DEGVRTYLMK-----YIYKNDIKREYLEEFASRPPLRW 226
Db      1274 TEKSNLYQTYLGYNEISDYLLRNVPQTSSAVIDSTVLRKKQDNQAVLPKETNNVQFNSVF 1333
Oy      227 TE----WGEDFVKKRKKEVLSLKKKCSCTL-----RNNGTSNKTCDDNENC--G 270
Db      1334 DSNNNVQOTFEFAKKLRKRISSKGRYRSLSLARYLTATRKPRVLGLDNLKIDITTLQ 1393
Oy      271 ACTQCEKY-----KKWEMRWKKHYSQ--KKKQLYKNSATY 306
Oy      1394 AFITKEEKQDSLNTLQKQDELITSLKSKQITKRSRHSKRRSHQFSRHHYTKRKH 1453
Oy      307 NNG-LAAVEANSEETYKNDPEVTEANSAKHARDYKLT--OLENMICTN 350
Db      1454 GNGGLRWYNNKKLKKFKKATNELRQWMMWSFLPRVYLSNQLVNNSTILT 1499

RESULT 11
SCPI_MESAU
ID SCPI_MESAU STANDARD; PRT; 845 AA.
AC 060563:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Synaptonemal complex protein 1 (SCP-1 protein) (Meiotic chromosome
DE synaptic protein) (Fragment).
GN SCPI OR SYN1.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036:
RX [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Testis;
RX MEDLINE=95181577; PubMed=7876343;
RA Dobson M.J., Pearlman R.E., Karatsakis A., Spyropoulos B.,
RA Moens P.B.;
RT "Synaptonemal complex proteins: occurrence, epitope mapping and
RT chromosome disjunction.";
RL Cell Sci. 107:2749-2760(1994).
CC -1- FUNCTION: MAJOR COMPONENT OF THE TRANSVERSE FILAMENTS OF
CC SYNAPTONEMAL COMPLEXES (SCS), FORMED BETWEEN HOMOLOGOUS
CC CHROMOSOMES DURING MEIOTIC PROPHASE (BY SIMILARITY). HAS NON-
CC SPECIFIC DNA BINDING CAPABILITY.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. IN TRIPARTITE SEGMENTS OF
CC SYNAPTONEMAL COMPLEXES, BETWEEN LATERAL ELEMENTS IN THE NUCLEUS.
CC FOUND ONLY WHERE THE CHROMOSOME CORES ARE SYNAPSED. ITS N-TERMINUS
CC IS FOUND TOWARDS THE CENTRE OF THE SYNAPTONEMAL COMPLEX WHILE THE
CC C-TERMINUS EXTENDS WELL INTO THE LATERAL DOMAIN OF THE
CC SYNAPTONEMAL COMPLEX.
CC -1- DOMAIN: CONSISTS OF AN ALPHA-HELICAL STRETCH OF 700 AA RESIDUES,
CC FLANKED BY N- AND C-TERMINAL GLOBULAR DOMAINS. THE C-TERMINAL
CC DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L32978; AAC42039.1;
DR Nuclear protein; Meiosis; Cell division; Phosphorylation;
KW DNA-binding; Coiled coil.
FT NON_TER 1
FT DOMAIN <1 672 COILED COIL (POTENTIAL).
FT 553 556 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 553 756 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

```

FT DOMAIN 830 838 ARC/LYS-RICH (BASIC).
 SQ SEQUENCE 845 AA; 99401 MW; D7F28873C824C6A8 CRC64;
 Query Match 5.7%; Score 109; DB 1; Length 845;
 Best Local Similarity 20.0%; Pred. No. 2.7;
 Matches 82; Conservative 57; Mismatches 179; Indels 92; Gaps 15;

QY 2 CKWOKLISFOIEKNNH-----NCKT-----EDAKKCKENT-----K 35
 D 36 CNLKETSARSARKTKKYEERETROYVDLNNNEKMLAEELRYOAEKRLDMHFK 95
 QY 36 LGEDEVCPMPRRONLCVHYLTKLNDCKEEDLREAFKSAAE-----TELLROYNS 89
 D 96 LKEDH-----EKIOHQEEYKKEVND--KENQVSLTIQTEKENKMKDLTFLEESRDK 148
 QY 90 KKNEDKILHRDMIPPEFFSMFTYTGDIYDCLDDIS---EKIADHYTTAKKITAY 146
 D 149 VNQLEKTKLQDENVELEKDKDLTSELEDTKMSIQRSNNQKALEEDQIATKTIYOL 208
 QY 147 FOKIGSKTTGKGVLEPEGWKEEYGLIMKGMCLALSYNETKKMGVATYLMKIYKN 206
 D 209 TE-----EKEAOMEFFNKAKTDHSHFVTELEKATCTLEELTEEQRLVKN 255
 QY 207 NDIKEYLEEFASPPPLRWVTEWGEDFVKNRKKEVLSLKK----- 246
 D 256 DQKILTMELQKSNLEDEWTK---FKNNNEVKLEELKTLAEDQKLDEKQOVERLAE 311
 QY 247 ----KDCSCLTRNNGSNKTCDDNENCGACKTCCEKTKKMEKMKHYSSQ-KKPOLYK 301
 D 312 ELQGGKDELLTLLOTRKEVHDLLEOLLVYKISDQNSKQVEBELKRLKEBEKLNALTA 371
 QY 302 N-SATYNNGLAVKEANS--ETYYKNDPEVTEANSKAHARDYLTQLEEM 346
 D 372 SCCKLSLNNKLTQFTNDMLALKYQEDIT--NSKKQEBMLK-QLENL 418

Db 372 SCCKLSLNNKLTQFTNDMLALKYQEDIT--NSKKQEBMLK-QLENL 418

RESULT 12
 ID RASO_THEAC STANDARD; PRT; 896 AA.
 AC 09HTR8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA double-strand break repair rad50 ATPase.
 GN RAD50 OR TA0157.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 OC Thermoplasmataceae; Thermoplasma.
 NCBI_TaxID=2303;
 [1]
 SEQUENCE FROM N.A.
 RX STRAIN-DSM 1728;
 RX MEDLINE=20479972; PubMed=110299001;
 RA Repp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermophilic scavenger Thermoplasma
 RT acidophilum.";
 RT accession: 407:508-513(2000).
 RL Nature 407:508-513(2000).
 CC - FUNCTION: Involved in DNA double-strand break repair (DSBR). The
 CC rad50/mre11 complex possesses single-strand endonuclease activity
 CC and ATP-dependent double-strand-specific exonuclease activity.
 CC Rad50 provides an ATP-dependent control of mre11 by unwinding
 CC and/or repositioning DNA ends into the mre11 active site (By
 CC similarity).
 CC - SUBUNIT: Forms a complex with mre11 (By similarity).
 CC - SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC EMBL: AL445063; CAC11304.1; -
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR003395; SMC_N.
 DR Pfam: PF02463; SMC_N; 1.
 KW DNA repair; Hydrolyase; ATP-binding; coiled coil; complete proteome.
 FT NE_BIND 30 37 ATP (BY SIMILARITY).
 FT DOMAIN 151 744 COILED COIL (POTENTIAL).
 SQ SEQUENCE 896 AA; 103435 MW; 2C584D700CCD2CB2 CRC64;
 Query Match 5.6%; Score 106.5; DB 1; Length 896;
 Best Local Similarity 20.5%; Pred. No. 4.2;
 Matches 94; Conservative 64; Mismatches 158; Indels 143; Gaps 22;

QY 10 SEQIEKNNHNCCKTTDADAKCKENTPLGE-DEVCVMPRRONLCVHYLTKLNDCKE-BD 67
 D 114 NDYIEKNI--TRSKDVELNSYFSKQGEDDLSGDPPARRK-----KLDELLEIEK 163
 QY 68 LREAF-----IKSAAETFLROYYN-----SKNVEDKILHRDM 102
 D 164 LEETIYVLKVDYDSLQAGISNLDYLISENERDRLKRYQDYVAELSKQIDEEAIESDL 223
 QY 103 IPPEFFRSMFTYTGDIYDCLDPTDISEKTIADHDVTTAK---KRTAVFOKIGSKT--- 154
 D 224 LRKKEASAEVNAVSKELIMLDATLKNMMSLSEANRYEERIKIDGLQELIGSTERYN 283
 QY 155 --INGKVLVER---GWMKEEYGLST-WKGMCLAL-----SYNETK----- 190
 D 284 EITSSVYVASRERIRYWDKQGIIDYRKRLKNDIGOVASYEDNMKKAELQADHDQYEI 343
 QY 191 ---MDE-----GVATYLMKIYTKNNNDI---KEYLEEFASPPPL----- 223
 D 344 MQRMOEIKHEDDLDTYESKYSLINELIQKKRKEEYRKQKDLGDELISRTLGAFAN 403
 QY 224 -----RWTEWGEDFVKNRKKEVLSLKKDCSCLTRNNGTSNKTCDNENCGA 271
 D 404 ASELVAIYEIRIDIDEINTD-LGNLQVIGALRQKEE--IRRN--NMMEGHKKCPV 457
 QY 272 CKTQCEKYYKWMERMKKHYSS-----QKKKPOLYKNSATYNNGLAV 312
 D 458 CGT--DLGDSRRIRREHYSIEDINLEITDILHERASALDEKRLISMESYLAKG--- 512
 QY 313 KEANSETYKN-----PPEVTEANSKAHARDYLTQLEEM 346
 D 513 KIREYETIDROMKDLAQITDDENSLSJYAKRTKYEQL 551

Db 513 KIREYETIDROMKDLAQITDDENSLSJYAKRTKYEQL 551

RESULT 13
 ID TOPL_CRIGR STANDARD; PRT; 767 AA.
 AC 007050;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA topoisomerase I (EC 5.99.1.2).
 GN TOPI OR TOP-1.
 OS Cricetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetulus.
 OX NCBI_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94064611; PubMed=8244980;
 RA Tanizawa A., Bertrand R., Kohlhaagen G., Tabuchi A., Jenkins J.,
 RA Pommer Y.;
 RT "Cloning of Chinese hamster DNA topoisomerase I cDNA and
 RT identification of a single point mutation responsible for
 RT camptothecin resistance";
 RL J. Biol. Chem. 268:25463-25468(1993).
 CC - FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE
 CC CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.

CC -1- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
 CC DNA, followed by passage and rejoining.
 CC -1- ENZYME REGULATION: SPECIFICALLY INHIBITED BY CAMPOTHECIN (CPT), A
 CC PLANT ALKALOID WITH ANTITUMOR ACTIVITY.
 CC -1- SUBUNIT: MONOMER.
 CC -1- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
 CC NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
 CC RELAX ONLY NEGATIVE SUPERCOILS.
 CC -1- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA
 CC BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN
 CC WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS
 CC AT ONE END OF THE ENZYME-SEVERED DNA STRAND.
 CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC TYPE I TOPOISOMERASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: Z21624; CAA/9747.1; -
 DR EMBL: Z21625; CAA/9748.1; -
 DR PIR: S32697; S32697.
 DR HSSP: P11387; 1A35.
 DR InterPro: IPR001631; Topoisomerase_I.
 DR Pfam: PF01028; Topoisomerase_I.
 DR Pfam: PF02919; Topoisomerase_I_N; 1.
 DR PRINTS: PR00416; EUTPISMRASE1.
 DR SMART: SM00435; TOPEUC; 1.
 DR PROSITE: PS00176; TOPOISOMERASE_I_EUK; 1.
 KM Isomerase; Topoisomerase; DNA-binding.
 FT ACT SITE 725 DNA CLEAVAGE (BY SIMILARITY).
 FT VARIANT 505 505 S -> G (IN CPT-RESISTANT CELL).
 SQ SEQUENCE 767 AA; 90867 MW; 347336D424EF35A9 CRC64;

Query Match 5.5%; Score 106; DB 1; Length 767;
 Best Local Similarity 20.5%; Pred. No. 3.8; Mismatches 106; Indels 114; Gaps 17;
 Matches 69; Conservative 47; Mismatches 106; Indels 114; Gaps 17;

QY 14 EKNINHNCKTDEDAKWCENTKLGDEGVCMPPRRONLCVHYLTRKLNDSKEEDLREAFI 73
 DB 89 EKREKEXIRASGDGAKIRKE-----KENGFSPPR-----IKDEPDDG----- 126
 QY 74 KSAAEFTLLROYYSKKNVEDDKILHRDMIPPEFRSMYTTGDRIDICLDI-DISE--- 129
 DB 127 -----YFAPRKEDIKPLKRPDEDD-----ADYKPKKIKTEDIKKEKK 164
 QY 130 -KIADHDVTTAKKITTAVFQIGSKTTNGKVLERE-----GMWKE-YGLSI-W----- 175
 DB 165 RLLEEDGCKLTKTKNNKDKKGAESDNKKKPKKEEDGKWKWWEERIPREIKKFLKH 224
 QY 176 KGMICALSYNETTKMKDEGVRTY---LMKYLYKNNDIKEYL-----EEFASRPFLR-W 225
 DB 225 KSPVAPAPY-----EPLEGEVRYFDGKVKLSPKAEVATFPKMLDHEYTTKEIFRKNF 280
 QY 226 VTEMEDFVKNKKKELVSLKKKCDSCITLRNNGTSNKTCDNENCGACQCKQCKYKKMMAR 285
 DB 281 FDMRKRKEMTNDKRNVTNL-SKCDP-----TOMSOYFRDOSE 316
 QY 286 MKKHYSSOKKPOLYKNSATYNNGLAVKENSETYK 321
 DB 317 ARKQMSKEK-----LKIKENEKLLK 338

RESULT 14
 RBPL_PLAVB STANDARD: PRT; 2869 AA.
 ID RBPL_PLAVB
 AC 000738;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Reticulocyte binding protein 1 precursor.
 GN RBPL.
 OS Plasmodium vivax (strain Belen).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 ON NCBI_TaxID=31273;
 RX MEDLINE=92315338; PubMed=1617731;
 RA Gallinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
 RT "A reticulocyte-binding protein complex of Plasmodium vivax
 RT merzoites."
 RL Cell 69:1213-1226(1992).
 CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
 CC HUMAN RETICULOCYTE CELLS.
 CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Membrane-bound.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: M88097; AAA29743.1; -
 KM Malaria: Receptor; Signal; Transmembrane.
 FT SIGNAL 1 17
 FT CHAIN 18 2869
 FT DOMAIN 18 2807 RETICULOCYTE BINDING PROTEIN 1.
 FT TRANSMEM 2808 2826 EXTRACELLULAR.
 FT DOMAIN 2827 2869 POTENTIAL.
 FT SITE 1030 1032 CYTOPLASMIC.
 FT SITE 2599 2601 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 2869 AA; 330213 MW; B9DBE44205EBCE CRC64;

Query Match 5.5%; Score 105.5; DB 1; Length 2869;
 Best Local Similarity 21.6%; Pred. No. 19; Mismatches 85; Conservative 47; Mismatches 123; Indels 139; Gaps 19;

QY 17 NIHNCK-TEDAKMKCENTKLGDEGVCMPPRRONLCVHYLT-----KL 59
 DB 998 SINCKKYNTDIDLRSKIKITLREEVQKEMPRKDCGENTALLKSLDRKMKINEKL 1057
 QY 60 ND-----DSKFEEDLEAFIKSAAEFTLLROYYSKKNVEDDKI-LHRMIPPEFRSMF 112
 DB 1058 NGRNLNSIDTKRED-----LLKEYSESQ-----KHLSDQKGPQ---DPL 1096
 QY 113 YTFGDRIDICLDITD-----ISEKIADHDVTTAKKITTAVFQIGSK-----TTNGKVV 160
 DB 1097 NKIDEMEDIKRVDDELNVAYQYISENKVYLFKNNSVTYIEAMHSHINTYVAHGITSKNEI 1156
 QY 161 LEREGWKKETGLSIWGMICALSYNETTKMKDEGVRTYLMKYLYKNNDIKEYLEEFASRP 220
 DB 1157 LKSV-----KEV-----EDKLNLEQUNEDYKVV-----KNPENKOLEAIRGSM 1195
 QY 221 PLRLVWTEGDEGFVKNRKKELYSLKKKCDSCITLRNNGTSNKTCDNENCGACQCKQ- 276
 DB 1196 SKLK-----EVINKHSEMTOLESTAN--FLKSNAKKEHDEHLELNTKTKGQWRDIY 1246
 QY 277 EYKKWMEKRRKKHYSSOKKPOLYKNSATYNNGLAVKENSETYKNDPE----- 325
 DB 1247 EKLKTKAEFLKE-----GTVNE-----LKDNANEKANRYEPPEPENNIIIGHYLE 1288
 QY 326 -----VTEANSAKHARDYIKTOLENNI 347
 DB 1289 RITVEKDKAGKYVEEWS-----LKTRIEKLII 1315

RESULT 15
 ENPL_CATRO STANDARD: PRT; 817 AA.
 ID ENPL_CATRO
 AC P35016;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 14:54:09 : Search time 54.8604 Seconds
(without alignments)
1318.301 Million cell updates/sec

Title: US-10-087-013-9
Perfect score: 1913
Sequence: 1 PCKMVKLISEQIEKNKIN.....AKHARDYKLTQENMCTING 351

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
1 number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 1913 | 100.0 | 1685 | 5 | Q9U4A2 |
| 2 | 646 | 33.8 | 921 | 5 | Q25989 |
| 3 | 629.5 | 32.9 | 2658 | 5 | Q8T5G0 |
| 4 | 626.5 | 32.7 | 2527 | 5 | Q95W83 |
| 5 | 625.5 | 32.7 | 3078 | 5 | Q26031 |
| 6 | 562.5 | 29.4 | 3026 | 5 | Q26030 |
| 7 | 550.5 | 28.8 | 3542 | 5 | Q9U5M2 |
| 8 | 547.5 | 28.6 | 1711 | 5 | Q9U5M2 |
| 9 | 520 | 27.2 | 2706 | 5 | Q9U5M2 |
| 10 | 501 | 26.2 | 3006 | 5 | Q26032 |
| 11 | 500 | 26.1 | 3287 | 5 | Q8T326 |
| 12 | 495.5 | 25.9 | 1615 | 5 | Q8T325 |
| 13 | 460.5 | 24.1 | 2710 | 5 | Q9XZB8 |
| 14 | 434.5 | 23.2 | 2212 | 5 | Q94657 |
| 15 | 434.5 | 22.7 | 2169 | 5 | Q97312 |
| 16 | 433.5 | 22.7 | 2042 | 5 | Q25766 |

| | | | | | |
|----|-------|------|------|---|--------|
| 17 | 432 | 22.6 | 2924 | 5 | Q25733 |
| 18 | 429.5 | 22.5 | 2209 | 5 | Q9U0G6 |
| 19 | 418 | 21.9 | 176 | 5 | Q9B1F0 |
| 20 | 403.5 | 21.1 | 2209 | 5 | Q97324 |
| 21 | 403.5 | 21.1 | 2647 | 5 | P90580 |
| 22 | 402 | 21.0 | 184 | 5 | Q9B3E6 |
| 23 | 399.5 | 20.9 | 2664 | 5 | Q26033 |
| 24 | 398.5 | 20.8 | 2197 | 5 | Q96296 |
| 25 | 397.5 | 20.8 | 2135 | 5 | Q61077 |
| 26 | 384.5 | 20.1 | 2163 | 5 | Q9NFB6 |
| 27 | 375 | 19.6 | 2228 | 5 | Q60991 |
| 28 | 374.5 | 19.6 | 2182 | 5 | Q26034 |
| 29 | 364 | 19.0 | 182 | 5 | Q00832 |
| 30 | 360.5 | 18.8 | 1729 | 5 | Q25734 |
| 31 | 355.5 | 18.6 | 2277 | 5 | Q9U0G5 |
| 32 | 353 | 18.5 | 699 | 5 | Q25987 |
| 33 | 350 | 18.3 | 2647 | 5 | Q9U4X0 |
| 34 | 345 | 18.0 | 162 | 5 | Q9B3E7 |
| 35 | 344 | 18.0 | 160 | 5 | Q9B3E9 |
| 36 | 336.5 | 17.6 | 195 | 5 | Q00831 |
| 37 | 333.5 | 17.4 | 616 | 5 | Q8T9N3 |
| 38 | 333.5 | 17.4 | 616 | 5 | Q8T9L8 |
| 39 | 333.5 | 17.4 | 1210 | 5 | Q96724 |
| 40 | 333.5 | 17.4 | 1210 | 5 | Q8W531 |
| 41 | 329 | 17.2 | 616 | 5 | Q27361 |
| 42 | 329 | 17.2 | 1475 | 5 | Q25842 |
| 43 | 325 | 17.0 | 616 | 5 | Q25739 |
| 44 | 325 | 17.0 | 616 | 5 | P90585 |
| 45 | 323 | 16.9 | 616 | 5 | Q27390 |

ALIGNMENTS

| | |
|--|--|
| RESULT 1 | |
| ID Q9U4A2 | PRELIMINARY; PRT; 1685 AA. |
| AC Q9U4A2; | |
| DT 01-MAY-2000 (TREMREL. 13, Created) | |
| DT 01-MAY-2000 (TREMREL. 13, Last sequence update) | |
| DT 01-MAR-2002 (TREMREL. 20, Last annotation update) | |
| DE Variant surface protein PfEMP1 (Fragment). | |
| GN VAR. | |
| OS Plasmodium falciparum. | |
| OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium. | |
| OX NCBI_TaxID=5833; | |
| RN (1) | |
| RP SEQUENCE FROM N.A. | |
| RC STRAIN=IT; | |
| RX MEDLINE=20144115; PubMed=10677532; | |
| RA Smith J.D., Craig A.G., Krieck N., Hudson-Taylor D., Kyes S., Fagen T., Pinches R., Baruch D.I., Newbold C.I., Miller L.H.; | |
| RT "Identification of a Plasmodium falciparum intercellular adhesion molecule-1 binding domain: A parasite adhesion trait implicated in cerebral malaria." | |
| RT Proc. Natl. Acad. Sci. U.S.A. 97:1766-1771(2000). | |
| RL EMBL: AF193424; AAF18980.1; | |
| DR InterPro: IPR000454; Eub.ATPase_Csub. | |
| DR InterPro: IPR004258; PFEMP. | |
| DR Pfam: PF03011; PFEMP. 1. | |
| FT PROSITE: PS00605; ATPASE_C; 1. | |
| FT NON_TER 1685 | |
| SO SEQUENCE 1685 AA; 193671 MW; D1FDA2666B0551E CRC64; | |
| Query Match | 100.0%; Score 1913; DB 5; Length 1685; |
| Best Local Similarity | 100.0%; Pred. No. 7,1e-124; |
| Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| QY 1 PCKMVKLISEQIEKNKINCKTEDAKKCKCENTKIGEDBGVAMPRRONLCVHYLTKLN 60 | |
| DB 1224 PCKMVKLISEQIEKNKINCKTEDAKKCKCENTKIGEDBGVAMPRRONLCVHYLTKLN 1283 | |
| QY 61 DSKKEEDLEAFIKSAAEFTLLRQYNSKNVEDDKILHDMIPPEFRSMFTFGDYRD 120 | |

```

Db 1284 DSKKEEDLEAFIKSAAEETFLRLROYNSKNVEDKILHRDIPPEFRSMETG DYRD 1343
QY 121 ICLDPTDISEKTIADHDVTAKKKTITAVFOKISGKTNGKAKVLEREGWMEKYGISTIKGMLC 180
Db 1344 ICLDPTDISEKTIADHDVTAKKKTITAVFOKISGKTNGKAKVLEREGWMEKYGISTIKGMLC 1403
QY 181 ALSTYETFKKDEGVRYTLKMYIKYKNDIKKEYLEEFASRPFLRWVTEMGEDFYVKNRKE 240
Db 1404 ALSTYETFKKDEGVRYTLKMYIKYKNDIKKEYLEEFASRPFLRWVTEMGEDFYVKNRKE 1463
QY 241 LVSLAKKCDSCITLRNNGTSNKTCDNENCGACKTQCEKYKWKMEKWKHHYSOKKKPOLY 300
Db 1464 LVSLAKKCDSCITLRNNGTSNKTCDNENCGACKTQCEKYKWKMEKWKHHYSOKKKPOLY 1523
QY 301 KNSATYNNGLAVKEANSFETYNKDEPVEVANSKAKIARADYLTQLENMICTNG 351
Db 1524 KNSATYNNGLAVKEANSFETYNKDEPVEVANSKAKIARADYLTQLENMICTNG 1574

```

UNIT 2

```

ID 025989 PRELIMINARY; PRT; 921 AA.
AC 025989;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE (clone pS1H) ORF (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DD2/NM;
RX MEDLINE=95350219; PubMed=7624377;
RA Peterson D.S., Miller L.H., Wellem T.E.;
RT "Isolation of multiple sequences from the Plasmodium falciparum genome
RT that encode conserved domains homologous to those in erythrocyte-
RT binding proteins."
RL Proc. Natl. Acad. Sci. U.S.A. 92:7100-7104(1995).
DR EMBL; L38454; AACC37240.1; -.
DR InterPro; IPR004258; PfEMP.
DR Pfam; PF03011; PfEMP; 1.
FT NON_TER 1
SQ SEQUENCE 921 AA; 104711 MW; D98BC07763628746 CRC64;

```

```

Query Match 33.88; Score 646; DB 5; Length 921;
Best Local Similarity 41.08; Pred. No. 1.3e-36;
Lches 150; Conservative 54; Mismatches 112; Indels 50; Gaps 16;

```

```

2 CKWQKLIISQIEKNNIHC-KKTED-AKMCENT-KIGDEBVCMPRRONICVHYLR 58
Db 396 CEIYAEKLKKNRTTYGECYRKETISEWTCDSKIKMGO-HGACIPRROKLCIHLER 454
QY 59 LINDSKEDLEAFIKSAAEETFLRLROY--NSKVED-DKILHRDIPPEFRSMFY 114
Db 455 IMNTNE--LKVAFIKCAAEETFLRLWQNYKKDKNGAEDLDEKLGIIIPEDRKQMFY 512
QY 115 FGDRICDPTDI-SKKIADHDVTAKKKTITAVFOKISGKTNGKAKVLEREGWMEKYGIS 173
Db 513 FADYRICDPTDISSSKKDTSSKGVKVCNTDVFYKISN--SIRYKSMWETNGPV 566
QY 174 IMKGMICALSYNT-----ET-KKDEGVRYTLKMYIKYKNDIKKEYLEEFASRPFLRW 225
Db 567 IMKGMICALSYDTSLNANVPETHKLTLEGNNF-EKVI-F-GSSSTTLKFSRPOFLRW 624
QY 226 VTEMGEDEFYVKNRKEKVLVSLKCKDCSTLRNNGTSNKTCDNENCGACKTQCEKYKWKME 285
Db 625 LTEMGEDEFYVKNRKEKVLVSLKCKDCSDVDGDC--KC--NGKVCACKQCKQYHSMIGI 678
QY 676 WKHHYSOKKKPOLYKNSATYNNGLAVKEANSFETYNKDEPVEVANSKAKIARADYLTQLEN 345
Db 679 WIDNYKKQKGRYTEV-----KIPLVKEDKDVKNSD--ARDYLTQLEN 721

```

```

QY 346 MICTNG 351
Db 722 MKCVNG 727

```

RESULT 3

```

ID 08T560 PRELIMINARY; PRT; 2658 AA.
AC 08T560;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Erythrocyte membrane protein-1 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JDP8;
RA Chattopadhyay R., Pillai C.R., Chitnis C.;
RT "Identification of a domain responsible for binding to intercellular
RT adhesion molecule-1 from a Plasmodium falciparum field isolate."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY028643; AAK49742.1; -.
FT NON_TER 2658
SQ SEQUENCE 2658 AA; 301326 MW; ABE33BC9167E7B7 CRC64;

```

```

Query Match 32.98; Score 629.5; DB 5; Length 2658;
Best Local Similarity 39.48; Pred. No. 6e-35;
Matches 146; Conservative 59; Mismatches 107; Indels 59; Gaps 17;

```

```

2 CKWQKLIISQIEKNNIHC-KKTEDA-KKCENTKLGDE-GVCMPPRRONICVH 54
Db 1307 CKTVNDITLSDPTKQYOCMAKIKINISYPMWTVNSKFENNEYGPMPPRRONICLY 1366
QY 55 YLRKNDKEDLEAFIKSAAEETFLRLROYNSKNVEDKILHRDIPPEFRSMFY 114
Db 1367 YLRKNDKEDLEAFIKSAAEETFLRLWQNYKKDKNGAEDLDEKLGIIIPEDRKQMFY 1426
QY 115 FGDRICDPTDISEKTIADHDVTAKKKTITAVFOKISGKTNGKAKVLEREGWMEKYGIS 174
Db 1427 YGDRICDPTDISEKTIADHDVTAKKKTITAVFOKISGKTNGKAKVLEREGWMEKYGIS 1476
QY 175 WKGMICALSYNT-----ET-KKDEGVRYTLKMYIKYKNDIKKEYLEEFASRPFLRW 228
Db 1477 WKGMICALSYDTSLNANVPETHKLTLEGNNF-EKVI-F-GSSSTTLKFSRPOFLRW 624
QY 229 WGEDEFYVKNRKEKVLVSLKCKDCSTLRNNGTSNKTCDNENCGACKTQCEKYKWKME 287
Db 1531 WGEDEFYVKNRKEKVLVSLKCKDCSDVDGDC--KC--NGKVCACKQCKQYHSMIGI 678
QY 288 KHHYSOKKKPOLYKNSATYNNGLAVKEANSFETYNKDEPVEVANSKAKIARADYLTQLEN 345
Db 1587 KHHYSOKKKPOLYKNSATYNNGLAVKEANSFETYNKDEPVEVANSKAKIARADYLTQLEN 1637
QY 332 AKHARDYLTQ 342
Db 1638 YSKAGYIKKE 1648

```

RESULT 4

```

ID 095W83 PRELIMINARY; PRT; 2527 AA.
AC 095W83;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Erythrocyte membrane protein 1 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

```



```

Query Match          32.7%; Score 625.5; DB 5; Length 3078;
Best Local Similarity 38.8%; Pred. No. 1.3e-34;
Matches 137; Conservative 61; Mismatches 126; Indels 29; Gaps 9;

QY      2 CKNVQLKLSIQEIEKKNNINHCKTIED-AKKCENTRGLGEDEGVCMPPRRLNCVHYL---T 57
       11::: :::: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1347 CKIVEKEITLEGKNGRTVTGECPEKSPESYPDMDCNNDISHDCAQMPRRCKLTLYIAHES 1406
QY      58 KLNDSKSPEEDLRERFIISSAAAEFTLLROYNSKRWEDDKILHRDRIPEFRFSRYTFGD 117
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1407 QTENIKTKDDNLKDCAFITAFAEFLLSMOYYKSKNDSSEAKILDRLIPQFLRSMMYTFFGD 1466
QY      118 YRDICLDPTDISEKLADVDVTJAKKRITAVFOKISKSTTNKGVLREBGMKEYSLGWK 177
       |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1467 YRDICLWTDIDSK--QNADVAKAKRKIGKFYSKSDSKEPSG---LSROEMWKNTNPDLWGK 1521
QY      178 MCALS-YNTEIRKMDGVRTYLKKIYIKNNDILEYELEFPASRPPLRWYTEMGDEPYKN 236
       |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1522 MLCALFYIVDTDDNRKRIKINDSYDVKNQGNGNPISLEEFAPKAQPOLRWMIMEGEEPCAE 1581
QY      237 RKRELVSLLKKCBDSCTLRNNGTSNKTKCDNENCGACATQCCKYIKRMWRMKKHHSQQKK 296
       |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1582 RÖKKENIETKDAC-----NEINSYQGCDAKH--RCNQACRAYOEYENKKKEFSGGQTNN 1633
QY      297 FOALKNSATYNAGLAVERKANSETYKNDPEVEANSAKHARDYLTOLNMICT 349
       |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      1634 FVLKANQPPOD-----PEYKYCEKDGVOPIQGN-----EYLQKCDNNKCS 1675

RESULT 6
AC ID 026030 PRELIMINARY; PRT: 3026 AA.
OC 026030:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Variant surface protein (Fragment).
GN VAR.
OS Plasmodium falciparum.
CC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_taxonomy:5833;
RN [1]
RP SEQUENCE OF 2044-2922 FROM N.A.
RA STRAIN-IT 4/25/5;
RC MEDLINE=95330813; PubMed=7606788;
RX Su X.Z., Heatswole V.M., Wertheimer S.P., Guinet F., Herrfeldt J.A., Peterson D.S., Ravetch J.A., Wellens T.E.;
RA "The large dense gene family var encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes.";
RT Cell 82:89-100(1995);
RL [2]
RN [2]
RP SEQUENCE OF 2044-2922 FROM N.A.
RA STRAIN-IT 4/25/5;
RC Smith J.D., Chitnis C.E., Craig A.G., Roberts D.J.,
RA Hudson-Taylor D.E., Peterson D.S., Pinches R., Newbold C.I.,
RA Miller L.H.;
RA "Switches in the expression of Plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes.";
RT Cell 0:0-0(1995);
RL [3]
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN-IT 4/25/5;
RC MEDLINE=99094502;
RX Smith J.D., Kyes S., Craig A.G., Fagan T., Hudson-Taylor D., Miller L.H., Baruch D.I., Newbold C.I.;
RA "Analysis of adhesive domains from the A4VAR Plasmodium falciparum erythrocyte membrane protein-1 identifies a CD36 binding domain.";
RL Mol. Biochem. Parasitol. 97:133-148(1998);
RN [4]
RP SEQUENCE FROM N.A.
```

RC STRAIN=IT 4/25/5;
 RA Smith J.;
 RA Submitted (Jul-1995) to the EMBL/GenBank/DBJ databases.
 RN (5)
 RP SEQUENCE FROM N.A.
 RC STRAIN=IT 4/25/5;
 RA Ryes S., Smith J.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: 142244; AAD0351.1;
 DR InterPro: IPR004258; PfEMP.
 DR Pfam: PF03011; PfEMP; 2.
 FT NON_TER 3026 3026
 SQ SEQUENCE 3026 AA; 348124 MW; EC0A971A82193788 CRC64;
 Query Match 29.4%; Score 562.5; DB 5; Length 3026;
 Best Local Similarity 34.0%; Pred. No. 3e-30;
 Matches 128; Conservative 53; Mismatches 133; Indels 63; Gaps 12;
 2 CMAVOKLISEQLEKNNHNC-----KTEDAKMKCENTKLGEDGVCMPRRONLCVHYT 55
 1987 CEIVGEILNGODGCTKKIECNKRYPTKNDYPGWNCIDKVINREBSCMPRROKLCIHN 2046
 56 LTKLNDSEKEDLREAFIKSAAEFTLLRQYNSK-----NVEDD-----KILHR 100
 2047 LEHLSEKATELETKRAFIKCAIEFTFWLMDRYKEDKKDEKTEGGGISDDPDPOKKEG 2106
 101 DMIPPEFFSMETTGDDYRDICLDIDISEKIDHDTTA-KKITAIVFQKISKTNGK 159
 2107 GTIPEDFKQMYTYGDRDLEFGDISK--GHEGKSALGKRKIDSLFKNGDOKSPSGKT 2163
 160 VLREBGWKEVGLSTIMKGMICALSYNTEKMKDEGVRTYLMKRYIKNN---DIKEYLEEF 216
 2164 PTE---WMNDYGDPLMKGVCGLSHH-----IKNGKEDLRKRLIDNNKRYTKISKLEDF 2215
 217 ASRPFLKRVATWGEDVFNKKEKELVSLKKCD--SCLIRNNGTSNKTCDDNENGCACKT 274
 2216 ASRPFLKRVATWGEDVFNKKEKELVSLKKCD--SCLIRNNGTSNKTCDDNENGCACKT 2265
 275 OCEYKKMKMERKWKHYSSQKKKELYKNSATYNNGLAVNEANSERYKNDPEYTEANSACK 334
 2266 ACEATKSMILKMDKDYEQDTAKFD-----KDKKDKFDTGSAEVDVAANSS 2311
 335 ARDYKLTQLENNICNG 351
 2312 VHEVLOEELKN-LCTKG 2327
 RESULT 7
 0905M2 PRELIMINARY; PRT: 3542 AA.
 01-MAY-2000 (TREMBlrel. 13, Created)
 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 FC83 CSA ligand (Fragment).
 GN VAR.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_Taxid=5833;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=FC83;
 RX MEDLINE=20006305; PubMed=10535993;
 RA Buffet P., Gamaun B., Scheidig C., Baruch B., Oishi S., Fujii N.,
 RA Fuesi T., Pary D., Miller L.H., Gysin J., Schertl A.,
 RT "Plasmodium falciparum domain mediating adhesion to chondroitin
 RT sulfate A: A receptor for human placental infection."
 RT Proc. Natl. Acad. Sci. U.S.A. 96:12743-12748(1999).
 DR EMBL: AJ13811; CAB59840.1;
 DR InterPro: IPR001219; Neurotoxin.
 DR InterPro: IPR004258; PfEMP.
 DR Pfam: PF03011; PfEMP; 1.
 DR PRINTS: PR00284; TOXIN.

FT NON_TER 3542 3542
 SQ SEQUENCE 3542 AA; 413089 MW; 970D85EE8DA2EC2 CRC64;
 Query Match 28.8%; Score 550.5; DB 5; Length 3542;
 Best Local Similarity 39.0%; Pred. No. 2.4e-29;
 Matches 133; Conservative 48; Mismatches 115; Indels 45; Gaps 14;
 2 CMAVOKLISEQLEKNNHNC-----KTEDA--KKCENTKLGEDGVCMPRRONLCVHYT 57
 1261 CKTVNDILKENDGKQVEDCHPRKNSNGYPMQCCNINIVEDPRVCMPRROKLCVHFLA 1320
 58 KLNDSK---EEDLREAFIKSAAEFTLLRQYNNKNVED---DRLHDMIPPEFFSM 111
 1321 NDNELKKLOSQYNLKEAFIKSAAEFTPSWYTKSKDGBGNLDELKGLIPPAFLSM 1380
 112 FYTFGDYRDICLDIDISEKIDHDTTAKKITAIVFQKISKTNGKRYLREGWMEYK 171
 1381 FYTFGDYRDLEFGDISK--GHEGSKLKEQIDSLFKNGDOKSPNGK--TQEWMTES 1435
 172 LSIKGMICALSYNTEKMKDEGVRTYLMKRYIKNN---NDIKEYLEFRASRPFLRWTE 228
 1436 HEIWEAMCAL-VKIGAKKD-----FENGYNNVKSFSKSTLEEFKRPQFLRWTE 1489
 229 WGEDVFNKKEKELVSLKKKDCSCTLRNNGTSNKTCDDNENCGA-CKTCEKMKMERK 287
 1490 WYDDCYTRQKLTAKVOEKCS-----NDOLKCTEKNKCEDYVYTKMK-K 1535
 288 KHYSSQKKKFFQLYKNSATYNN---GLAYKE--ANSEYKN 322
 1536 KEMIPQDKYYKDERDKRRFRQHGIGVWVTDYGTNATDYLN 1576
 RESULT 8
 096108 PRELIMINARY; PRT: 1711 AA.
 AC 096108;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PfEMP1.
 GN PF0010W.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_Taxid=5833;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99021743; PubMed=9804551;
 RA Gardner M.J., Tetteh H., Carucci D.J., Cummings L.M., Aravind L.,
 RA Koonin E.V., Shallow S., Mason T., Yu K., Fujii C., Pederson J.,
 RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pettea M.,
 RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
 RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
 RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
 RT falciparum."
 RT Science 282:1126-1132(1998).
 DR EMBL: AE001366; AAC71792.1;
 DR InterPro: IPR004258; PfEMP.
 DR Pfam: PF03011; PfEMP; 1.
 SQ SEQUENCE 1711 AA; 196550 MW; 7BDBE652742235B CRC64;
 Query Match 28.6%; Score 547.5; DB 5; Length 1711;
 Best Local Similarity 37.3%; Pred. No. 1.7e-29;
 Matches 128; Conservative 53; Mismatches 105; Indels 57; Gaps 14;
 29 WKC---ENWTKGDEGVCMPRRONLCVHYLTKLNDSEKEDLREAFIKSAAEFTLLRQ 85
 86 WMCINPSTGTSKDCADICVPRTOELCYLYKELS-DTQKGLRAFIKTAAGETIYLLMO 914
 86 YNSK-----NVEDDKI-LHRDMIPPEFFSMETTGDDYRDICLDIDISEKIDH 134
 915 KYKEDQNETASTEIDIDIDPQTOLNGSEIPEDFKROMETFGDYRDLEFGHYIG-----N 969
 135 DVTAKKKTITAVFQKISKTNGKVV-LERBGWKEVGLSTIMKGMICALSYNTEKMKDE 193

Db 1037 YMRKKKEEWNLQDKYYKDKRENKGIDKGPIGIIVKDYVLANAKEYLKKFTASCVTSSGK 1096

DE PFEMPI (Fragment).

| | JUL 12 | | |
|--------|--|--------------|-------------------------|
| 08T325 | | | |
| ID | 08T325 | PRELIMINARY; | PRT; 1615 AA. |
| AC | 08T325; | | |
| DT | 01-JUN-2002 (TREMBrel) | 21, | Created) |
| DT | 01-JUN-2002 (TREMBrel) | 21, | Last sequence update) |
| DT | 01-JUN-2002 (TREMBrel) | 21, | Last annotation update) |
| DE | PFEM1 (Fragment). | | |
| GN | TM284VAR3. | | |
| OS | Plasmodium falciparum. | | |
| OC | Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium. | | |
| NCBI | NCBI_TaxID=5833; | | |
| RNA | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=TM284; | | |
| RR | MEDLINE=21927235; PubMed=11930336; | | |
| RT | Rowe J.A., Kyes S.A., Rogerson S.J., Babiker H.A., Raza A.; | | |
| RT | "Identification of a conserved Plasmodium var gene | | |
| RT | implicated in malaria in pregnancy." | | |
| RT | J. Infect. Dis. 185:1207-1211(2002). | | |
| EMBL | AJ420412; CAD20866.1; | | |
| FT | NCBI_TER 1615 | | |
| ISO | SEQUENCE 1615 AA; 188639 MM; FCA896C00866DEA6 CRC64; | | |

[illegible]

```

Db      1019 IFPGTJSSCRATKIDTSQTKSLGDOATTEKGDTHIDDKKLOE---WMTIHGPKIMEG 1075
QY      178 MICALSYN--TEFKKMEGRTYLMKYYK--NNDIIE--YLEEFASRPPLRAVTEMGED 232
      1076 MICALTNGISESEK-----KNILQDYSYNKLNAEKDDCCLEKFAKPOFLRWYEMSD 1130
QY      233 EVKRRKEVLKSLKCKDSCRLRNNGTSNKTCDNENCGACKTCCEKYYKMKMEKKHYSS 292
      1131 FGEERKK---LEDKVEDVCIK--AKDYECGRKNKNSNCKVKKCKEYENITGKTKQYES 1184
QY      293 QKKKPOLYKNSATYNNGLAVKEANSETYKNDPEVTANSAKHARDYLTQLENNMICTG 351
      1185 QEGKF-----NTERKORCKPEY--NSYSKKDASEYLKDK-----CIDG 1219

RESULT 14
094657 PRELIMINARY; PRT; 2212 AA.
094657
01-FEB-1997 (TEMBLrel. 02, Created)
01-FEB-1997 (TEMBLrel. 02, Last sequence update)
01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Erythrocyte membrane protein 1.
GN PEMPL.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JG;
RA Yang S.Q., Wollish W.S., Gut J., Wu J., Ahn J., Petersen C.,
  Yang S.Q., Aikawa M., Leach J.H., Nelson R.G.;
  "The molecular cloning and DNA sequence analysis of Plasmodium
  falciparum erythrocyte membrane protein 1."
RT falciparum erythrocyte membrane protein 1.
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U31083; AAB06961.1;
  InterPro: IPR004258; PFEMP.
DR Pfam: PFO3011; PFEMP. 2.
SQ SEQUENCE 2212 AA; 253594 MW; A33456648C652FC CRC64;

Query Match 23.2%; Score 443; DB 5; Length 2212;
Best local Similarity 26.2%; Pred. No. 3.8e-22;
Matches 128; Conservative 59; Mismatches 133; Indels 168; Gaps 19;

QY      8 LISQLEKNNIHCKTE-----DAKWK-----ENTKGED----- 39
      1272 IVAGVLTKELENACPTKYGPKAPTSMKCIPTKTAATGSESSGNGALORAKRAVES 1331
QY      40 -----EGCMPRRONLCVHYLTKLND-----DSKEE 66
      1332 GSPVTSNGSGICIPRRRL--YIOKLHWAGNIVVSGAQTPOGCTSSPSGKEPSPD 1388
QY      67 DLREAFIKSAAEFTLL--ROYNSKVE-----DD 95
      1389 KLTAFIQAIAETFFLMRYKKEKEKEKRVANGGLVPSLNGPQPGVGTGDSPOS 1448
QY      96 KILHRMIPPEFRSMFYTGDRDCL---DIDISKIDHDTYTRAKKITAVPOKI-- 150
      1449 KLOQGTGVLPPPLRQMFYTLGDYADIFGKNDIVIDTKNGDKDIAEREKKIKADIEVLK 1508
QY      151 --GSKTTNGKVLREGMKVEGLSIKMGICALSYNTEK-----KMDGVRTYL--- 199
      1509 NADSQPPSPDK---RQTMWQNGEHTINGMICALTYKEKDEKGTPLKONGGLSALMDEK 1565
QY      200 -----MKYY-----KND--IKYLEEFASRPPLRAVTEMGEDFYKNR 238
      1566 NKRPKDOKYOYDKVLKEDENSGTSPKTNHVPPLPLNFISRPYFLREMGETFCERK 1625
QY      239 KELVSLKCKDSCRLRNNGTSNKTCDNEN-----CGACKTCCEKYYK 282
      1626 KRLKTRKVEK---WMDGKKOKSGDGEDCEIRKQDYSTVDFYCPPEGKYCRFYKRW 1681

```

```

QY      283 MERKKHYSSOKKPOLYKNSATYNNGLAVKEANSETYKNDPEVTANSAKHARDYLTQ 342
      1682 IEKKKDEYDKOKKQKAYNNQKTDARRN-----NDNAFS-----TTLDTCTAGDFLOT- 1728
QY      343 LENNMICTN 350
      1729 LKNGPCKN 1736

RESULT 15
097312 PRELIMINARY; PRT; 2169 AA.
097312
01-MAY-1999 (TEMBLrel. 10, Created)
01-MAY-1999 (TEMBLrel. 10, Last sequence update)
01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Var (3D7-varT3-2).
GN PFC1120C, MAL3P7.55.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=99376085; Pubmed=10448855;
RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
  Churcher C.M., Craig A., Davies R.M., Devlin K., Felkwell T.,
  Gentles S., Gwilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
  Horrocks P., Jagels K., Jassal B., Kyes S., McLean J., Moule S.,
  Ruggall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
  Rutter S., Skellern J., Squares R., Squares S., Sulston J.E.,
  Whitehead S., Woodward J.R., Newbold C., Barrell B.G.;
  "The complete nucleotide sequence of chromosome 3 of Plasmodium
  falciparum."
RL Nature 400:532-538(1999).
DR EMBL; AL034559; CAB39061.1;
  InterPro: IPR004258; PFEMP.
DR Pfam: PFO3011; PFEMP. 2.
SQ SEQUENCE 2169 AA; 246913 MW; 1D48ACB7AFAE3BE CRC64;

Query Match 22.7%; Score 434.5; DB 5; Length 2169;
Best local Similarity 26.7%; Pred. No. 1.4e-21;
Matches 128; Conservative 50; Mismatches 127; Indels 175; Gaps 16;

QY      1 PCKMYQKLSQLEKNNIHNC-----KTEDAKWK-----ENTLGEDEG----- 41
      827 PCDIVQTLF--ESTKMLSDACGLKYGGEKKEPPMKCISSGSDTGSTTKONDSGSEGG 884
QY      42 -----VCMPPRRONLCVHYLTKLNDKSEED----- 67
      885 HREKRTHESSDSTTSSGVCVPRRRKLYVGETLTKAEARKSSSTSPQEGSGVANASA 944
QY      68 -----LREAFIKSAAEFTLLROYNSKVEDKI----- 97
      945 SSTSPDATQTLRLDAFIQSAIAETFLMRYKAEENTRONKSPGLNGVAQFGSGYSGSE 1004
QY      98 -----LHRMIPPEFRSMFYTGDRDCL---DIDISKIDHDTYTRAKKITAVPOKI-- 140
      1005 SEKTPQOQWQSGTTPDPLRQMFYTLGDYADIFGKNDIVIDTKNGDKDIAEREKKIKADIEVLK 1064
QY      141 --KKTAVFQKIG--SKTNGKVLERE--GMMKEGLSIKMGICALSYNTEK-----K 189
      1065 ISERKIKVIEKSGDTPSRTPGQPSNDKSMWKHLAPITWGMITVALTYLDTINTASGEK 1124
QY      190 KMDGVRTY-----LMKYYKNDIK----- 211
      1125 KIEKDAVYKRLKMDAANNKPKKNGOODYTYEYKVEIKEDSGOKASTASQTPSPRASGEN 1184
QY      212 ---YLEEFASRPPLRAVTEMGEDFYKNRKKELVSLKCK--DSCRLRNNGTSNKTCD 265
      1185 KPTLLQSEVAKRPYFLREMGETFCERKKRLEKIKVDVEENGTARGGTQKQKSGD 1244
QY      266 NENCG-----ACKTCCEKYYKMKMEKKHYSSOKKPOLYKNSATYNNG 309

```

Db 1245 GEECSKIYENKDKIFKDLKPSKATPCGLYKRWIRKKDEYNKOKSAVNEQTK--YENG 1302

Search completed: June 20, 2003, 15:05:21
Job time : 56.8604 secs